Íris Cristina da Luz Batalha

Mestre em Biotecnologia

Engineered structures for the profiling and enrichment of the phosphoproteome

Dissertação para obtenção do Grau de Doutor em Bioengenharia (MIT-Portugal)

Orientador: Prof. Doutora Ana Cecília Afonso Roque, FCT-UNL
Co-orientador: Prof. Doutor Christopher Robin Lowe, University of Cambridge
Co-orientador: Doutora Olga Iranzo Casanova, ITQB-UNL

Júri:

Presidente: Prof. Doutora Maria Ascensão Carvalho Fernandes Miranda Reis
Arguentes: Doutor Abid Hussain
Prof. Doutora Ana Cristina Mendes Dias Cabral

Vogais: Prof. Doutora Ana Cecília Afonso Roque
Doutora Ana Sofia de Sousa Valente Coroadinha
Prof. Doutor Christopher Robin Lowe
Prof. Doutor Carlos Miguel Calisto Baleizão

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Julho 2014

ENGINEERED STRUCTURES FOR THE PROFILING AND ENRICHMENT OF THE PHOSPHOPROTEOME

To Susana Marques, Isabel Grilo, José da Luz, Luzia da Luz, and Valério Batalha

"O valor das coisas não está no tempo que elas duram, mas na intensidade com que acontecem. Por isso, existem momentos inesquecíveis, coisas inexplicáveis e pessoas incomparáveis."

"The value of things is not the time they last, but the intensity with which they occur. So there are unforgettable moments, inexplicable things and incomparable people."

Fernando Pessoa

ACKNOWLEDGEMENTS

This section of my thesis is probably the one that took me longer to start writing up. Not because I don't know what to say, but because I have so many things to share and people to share with that it is difficult to keep it short and simple (which I don't think I will be able to do). Four years don't seem much, but sometimes it feels like a lifetime. There is a sentence I like from Simone de Beauvoir that says "One's life has value so long as one attributes value to the life of others, by means of love, friendship, and compassion".

I will start by thanking to my supervisor, Cecília, for all these years of friendship and support. I will always remember fondly the moments I spent in "lab dos órsinhos", working in the lab, our trips, our group days, our dinners, and even our arguments. Thank you for being there for me.
I would also like to thank to my co-supervisors - Chris and Olga. Chris, thank you for accepting me in your lab in Cambridge, for the fruitful discussions, for being straightforward and always telling me what you think, and for all the funny moments. Olga, thank you for showing me a new world of peptide synthesis, for your support in the lab, and for being a righteous person.
The three of you made me better in many ways! I truthfully cherish you all, and I consider myself a lucky person for having you as my supervisors.

During the time I spent in Cambridge, I was also working in the Cambridge Centre for Proteomics. I would like to express my gratitude to Dr. Kathryn Lilley and Dr. Houjiang Zhou for giving me the opportunity of being there and for their help with the proteomic experiments.

I have a very special person I want to thank, Ana Pina, for being one of the best people I have ever met, for her giant heart, for being always, always, always there for me. I want to thank you for staying late with me in the lab, for hearing me complain whenever I need to, for taking me to the hospital when I have my migraines, for spending two hours per day with me on the phone when we are apart. Thank you for being such a great friend. I know in my heart that you will stay forever.
I want to thank to my other friends from the lab (I refuse to call you colleagues): Margarida, for the great moments we have spent together. I had a great time when you visited me in Cambridge. Telma, for making me laugh all the time and for your friendship. Sus, for the funny moments we've shared in the lab and for being my "attention-to-detail" sister. Vijay, for teaching me Marathi and for our amusing talks in the lab. Ricardo, for being the modelling master and helping me whenever needed. Henrique, for making everything funnier with your dark humour. I can never forget a few past members of our lab: Abid, for your friendship and for transforming every situation into a joyful moment. Sara and Raquel, for your friendship. Matheus, for being my first "student" and becoming such a good friend. Bá, for the time we spent together in the lab and during the summer, and for making everything seem better with your smile.

I would never forget the administrative staff: Maria José Carapinha, Maria Isabel Rodrigues, and Maria do Carmo Rodrigues for their assistance in countless situations. I would also like to acknowledge Maria da Palma Afonso and Maria Mafalda Manita, for their help in the lab. I want to express my appreciation to José Silva Lopes for all the administrative work in the MIT-Portugal program, and for his friendship. My acknowledgements to Carla Rodrigues for the ICP-AES analysis.

I would also like to thank a few people from Cambridge: Graziella, Rhian, Chen, Shrishti, Mohamed, Ke Xu, Shaleem, Colin, Jamie, Anil, Fernando, Ali, Nan, and Safwan. I want to express my deep gratitude to Jacky and Isik, who gave me full support during my stay in Cambridge, and to Kheng and Basma, for the funny moments and our marvellous dinners. I miss you all. I would like to show my appreciation to the assistant staff: A. Bishop, C. Jefferies, F. Hall, L. Hares, C. Heinrich, L. Hurst, D. Singh, and J. Walker; and to Sue Donelan from Selwyn College.
From Olga's lab, I would like to thank to Ana Fragoso, who helped me countless times in the lab, and also Anaís, Sara, Vanessa, Helder, Sílvia, and Rui. I would like to thank to the Mass Spectrometry

Laboratory at the Analytical Services Unit of the Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa for the mass spectrometry analysis.

A special thanks to Dr. David Rueda for accepting me in his lab at the Wayne State University and for the opportunity he gave me of learning a completely different scientific field. I enjoyed a lot the time I spent there. My deepest gratitude to my friend Elvin, who taught me everything in the lab, and for being always so cheerful. You are one of the best people I have ever worked with. I also want to thank to Gayan and Rajan for helping me when I first arrived to the lab, and to all the other lab members for making me feel so welcome; to Ran Ran and Joel; to my housemates: Miu, Allison, and Ethan, for all the funny moments we´ve shared; and to Mr. Guyer for receiving me in his house. Lucia, my sweet friend, I will never forget our road trip and the months we spent together in Detroit.

I would like to thank Prof. João Paulo Crespo and Prof. Patrícia Rosado Pinto for giving me the opportunity to join the Doctoral School committee of PhD students, and to all my colleagues in that committee.

To my mother, Fátima: You are the most important person to me in the world. I love you more than anything. You are an iron woman and I am so proud of being your daughter.

Guida, I am who I am today in large part because of the way you raised me. You're my family.

To my aunt, Angélica, and uncle, Josué, I will never have enough words to thank everything you have been doing for me during my entire life. I would never be here if it wasn't for you. I love you so much.

I am so lucky of being part of a family of strong women like my granny, Lucília! You are my rock star! My heart feels full of joy whenever I am around you.

I want to thank my father - João, my sisters – Sandra and Rita, and Vivelinda for their love and support.

Teresa, my "big sister" and Ricardo Sampaio, thank you for your friendship and for all the nice moments we've shared together. Mariana, my "old" friend, I am so fortunate to have you in my life. You're my friend, my family, and my confidant. To my "brothers", João Zorro and João Pedro Freire, I love you guys. I can always count on you, even when you're far. I trust you with my life. Ana Esteves, you are a remarkable woman, best friend ever, Babe! Mara, it doesn't matter for how long we stay apart, you will always be a huge friend! I miss you. Lucília Galha, thanks for being such a dear friend, for never forgetting me, and for watching the same TV series that I do. Rita and Mónica, thank you for being the best MITS ever! I love you girls! Mário, my "coffee break" friend, thank you for your friendship and all the nice moments we have spent together.

I also want to thank to my cousin Rodolfo, Inês Grilo, Nuno, Saruga, Molefas, Sónia Dias, Soraia, Mena, João Freire, Manuel, Rosa, Janeca, Jô, Bete, Guidinha Ramos, Luísa, Esmeralda, Renata, Glória, Inês Mexia, Zé Bulhão, J. Grilo, Abel, Miguel, R. Nascimento, R. Barros, Ana Torres, Sobral, Marta, Zé Rui, João Fialho, João Dias, Ana Almeida, Rimenys, Abby, and Michaela, for their friendship and support. Lena, Jorge, and João Carreira, thank you for your support and affection.

I want to thank Ricardo, to whom I have shared eight years of my life, for his unconditional love and friendship. We had ups and downs, but we always have made through the dark moments, we've always found each other in the sun.

I would like to thank the financial support from Fundação para a Ciência e a Tecnologia, Ministério da Educação e da Ciência, Portugal, for my doctoral fellowship (SFRH/BD/64427/2009) and the project PTDC/EBB-BIO/102163/2008 assigned to Prof. Cecília Roque, and also to the Associate Laboratory REQUIMTE (PEst-C/EQB/LA0006/2013).

ABSTRACT

Protein phosphorylation regulates a variety of complex cellular signalling networks. Its monitoring is fundamental for understanding the origins of various diseases and drive the development of novel therapeutics.

This work focused on the development of peptidomimetic ligands for the selective binding of phosphorylated peptides. Structural studies of the natural human phosphoprotein-binding domains (PBDs) revealed the existence of key amino acid residues involved in phosphate recognition: Arg, Lys, Tyr, Ser, His, Gly, and Asn. Based on this information, three solid-phase combinatorial libraries (232 small synthetic affinity ligands in total) were synthesized and screened against target phosphorylated peptides. The synthesis of the libraries was based on different chemical reactions: nucleophilic substitutions on a triazine ring, Ugi multicomponent reaction, and tandem Petasis-Ugi multicomponent reaction. Two lead ligands which selectively bound and eluted phosphorylated peptides were identified. The lead ligands were A8A3 and A8C2 – based on the triazine and Ugi reactions, respectively - and both possess substituents mimicking the His residue. Ligand A8A3 was immobilized onto magnetic nanoparticles coated with silica and dextran and efficiently enriched a mixture of trypsin-digested α-casein, β-casein, and bovine serum albumin in mono- and multi-phosphorylated peptides, comparable to the results obtained when employing a conventional protocol with Ti^{4+}-IMAC.

In parallel, cyclic β-hairpin scaffolds mimicking the Brca1 C-terminal domain were rationally designed and synthesized by standard Fmoc solid-phase chemistry. The lead candidate, Cyclic-M2 with the sequence cyclo(Glu-Gly-Phe-Gly-Dap-Gly-Dap-^{D}Pro-Pro-Gly-Val-Arg-Thr-Gly) was successfully immobilized onto a chromatographic support (32 μmol/g). The resultant adsorbent was selective in binding to the phosphorylated target peptide allowing a 65% recovery of the phosphopeptide during elution.

Keywords: phosphopeptide enrichment, affinity ligands, peptidomimetics, multicomponent reactions, β-hairpin.

RESUMO

A fosforilação de proteínas regula inúmeros mecanismos de sinalização celular e, como tal, a sua monitorização é fundamental para compreender as origens de uma panóplia de doenças e promover o desenvolvimento de novos agentes terapêuticos.

Este trabalho focou-se no desenvolvimento de novas estruturas biomiméticas para a ligação selectiva de péptidos fosforilados.

Estudos estruturais dos domínios humanos de ligação a fosfoproteínas revelaram a existência de aminoácidos-chave envolvidos no reconhecimento do grupo fosfato: Arg, Lys, Tyr, Ser, His, Gly, and Asn. Com base nesta informação, três bibliotecas combinatoriais compostas por 232 ligandos sintéticos de afinidade foram sintetizadas e avaliadas quanto à capacidade de ligação a péptidos fosforilados. Estas bibliotecas basearam-se em diferentes reacções químicas: substituição nucleofílica num anel de triazina, reacção multicomponente de Ugi, e reacção multicomponente de Petasis-Ugi. Foram identificados dois ligandos principais, um baseado na reacção de triazina – A8A3, e outro baseado na reacção de Ugi – A8C2. Ambos apresentam componentes que mimetizam o resíduo de His na sua estrutura, ligam selectivamente a péptidos fosforilados a pH 3 e permitem a eluição eficiente das espécies fosforiladas.

O ligando A8A3 foi imobilizado em nanopartículas magnéticas revestidas com sílica e dextrano e enriqueceu eficazmente em fosfopéptidos mono- e multi-fosforilados uma mistura de α-caseína, β-caseína, e albumina de soro bovino (BSA) digeridas com tripsina.

Em paralelo, estruturas cíclicas de *β-hairpins* foram modeladas de forma a mimetizar o domínio Brca1 C-terminal e sintetizadas por química em fase sólida. A melhor estrutura, M2-cíclico, com a sequência ciclo(Glu-Gly-Phe-Gly-Dap-Gly-Dap-^{D}Pro-Pro-Gly-Val-Arg-Thr-Gly), foi imobilizada num suporte cromatográfico (32 μmol/g). O adsorvente contendo o M2-cíclico ligou-se ao péptido fosforilado com elevada selectividade, recuperando-se 65% do péptido aquando da eluição.

Palavras-chave: enriquecimento de fosfopéptidos, ligandos de afinidade, peptidomiméticos, reacções multicomponentes, β-hairpin.

TABLE OF CONTENTS

INDEX OF FIGURES

INDEX OF TABLES

ABBREVIATIONS

2,5-DHB – 2,5-Dihydroxybenzoic acid

A1 – N,N-Dimethyl-4,4'-azodianiline

A10 – 2-(Ethylthio)ethylamine

A11 – m-Xylylenediamine

A12/ Tris – Tris(hydroxymethyl)aminomethane

A2 – Amino-2-propanol

A3 – Phenethylamine

A4 – Tyramine

A5 – Cadaverine dihydrochloride

A6 – Tryptamine

A7 – Agmatine sulfate salt

A8 – Histamine

A9 – Isopentylamine

ACN – Acetonitrile

AP – 1-Pyrenemethylamine hydrochloride

APTES – 3-aminopropyltriethoxy silane

B1 – 4-Imidazolecarboxaldehyde

B2 – Propionaldehyde

B3 – Phenylacetaldehyde

B4 – 2-Ethylhexanal

B5 – 3-(Methylthio)propionaldehyde

B6 – Tetrahydrofuran-3-carboxaldehyde

B7 – Indole-3-carboxaldehyde

BC – Binding capacity

BCA – Bicinchoninic acid kit for protein determination

BMP – Bacterial magnetic nanoparticles

BRCT – Brca1 C-terminal

BSA – Bovine serum albumin

C1 – 4-Imidazolecarboxylic acid

C18 – Octadecyltrimethoxysilane

C2 – Conserved domain 2

C2 – Succinamic acid

C3 – 3-(4-Hydroxyphenyl)propionic acid

C4 – γ-Aminobutyric acid

C5 – 3-(Ethylthio)propanoic acid

C6 – Coumarin-3-carboxylic acid H

C7 – 3-Indoleacetic acid

C_{eq} – Concentration of peptide in solution at the equilibrium

CTD peptide - Peptide of the carboxy-terminal domain of RNA polymerase II

Cyclic-M0 – Cyclo(Cys-Gly-Gly-Gly-Gly-Gly-Gly-^{D}Pro-Pro-Gly-Val-Gly-Thr-Gly)

Cyclic-M1 – Cyclo(Cys-Gly-Phe-Gly-Ile-Gly-Ile-^{D}Pro-Pro-Arg-Val-Gly-Thr-Gly)

Cyclic-M2 – Cyclo(Glu-Gly-Phe-Gly-Dap-Gly-Dap-^{D}Pro-Pro-Gly-Val-Arg-Thr-Gly)

DAHC – Diammonium hydrogen citrate

Dap - Diaminopropionic acid

DCM – Dichloromethane

dd water – Distilled deionised water

DIEA – N,N-diisopropylethylamine

DLS - Dynamic light scattering

DMF - Dimethylformamide

EDC – N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide

EDT – 1,2-Ethanedithiol

EDTA – Ethylenediaminetetraacetic acid

EDTA-Na – Ethylenediaminetetraacetic acid tetrasodium salt hydrate

Ellman's reagent – 5,5'-Dithiobis(2-nitrobenzoic acid)

ESI - Electrospray ionization

FAT – Fluorescent affinity tag

$Fe_3O_4@Al_2O_3$ – Alumina-coated iron oxide magnetic particles

Fe_3O_4@C: Carbon-coated iron oxide nanoparticles

$Fe_3O_4@hYPO_4$ – Iron oxide core with an yttrium phosphate hollow porous affinity shell

$Fe_3O_4@Ti\text{-}mSiO_2$ – Titanium grafted magnetic mesoporous silica

$Fe_3O_4@ZrO_2$ – Zirconia-coated iron oxide magnetic nanoparticles

FHA – Forkhead-associated

Gec – Guanidinoethylcysteine

GET – Guanidinoethanethiol

GH – Gly-His dipeptide

GK14 – Gly-Ala-Ala-Tyr-Asp-Ile-Ser-Gln-Val-Phe-Pro-Phe-Ala-Lys

GK14-P – Gly-Ala-Ala-Tyr-Asp-Ile-pSer-Gln-Val-Phe-Pro-Phe-Ala-Lys

GLYMO–IDA – 3-Glycidoxypropyltrimethoxysilane with iminodiacetic acid

GMA - Glycidyl methacrylate

HATU – O-(7-azabenzotriazole-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate

HBTU – 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate

HD – Hydrodynamic diameter

HEPES – 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid

HPLC – High-performance liquid chromatography

HTS – High-throughput screening

IAA – Iodoacetamide

ICP-AES – Inductively Coupled Plasma - Atomic Emission Spectroscopy

IDA - Imidodiacetic acid

IMAC – Immobilized Metal Affinity Chromatography

k' – Capacity factor

K_A – Association equilibrium constant

K_D – Dissociation equilibrium constant

LC – Liquid chromatography

M0 – Cys-Gly-Gly-Gly-Gly-Gly-Gly-^{D}Pro-Pro-Gly-Val-Gly-Thr-Gly

M1 – Cys-Gly-Phe-Gly-Ile-Gly-Ile-^{D}Pro-Pro-Arg-Val-Gly-Thr-Gly

M2 – Glu-Gly-Phe-Gly-Dap-Gly-Dap-^{D}Pro-Pro-Gly-Val-Arg-Thr-Gly

MALDI-TOF – Matrix-assisted laser desorption/ionization time-of-flight

MeOH – Methanol

MES – 2-(N-Morpholino)ethanesulfonic acid

MH1 – Mad-homology 1

MH2 – Mad-homology 2

MNPs – Magnetic nanoparticles

MNP-Si-Si-Dex – Magnetic nanoparticles coated with silica and dextran

MNP-Si-Si-Dex-A8A3 – MNP-Si-Si-Dex functionalized with ligand A8A3

MNP-Si-Si-Dex-A8C2 – MNP-Si-Si-Dex functionalized with ligand A8C2

MOAC – Metal Oxide Affinity Chromatography

MS – Mass spectrometry

NHS – N-Hydroxysuccinimide

NMP – N-methyl-2-pyrrolidone

NMR – Nuclear magnetic resonance

NTA - Nitriloacetic acid

P(HEMA-GMA) – Poly(2-hydroxyethyl methacrylate-co-glycidyl methacrylate)

PAA – Poly(acrylic acid)

PA-Fe3O4@YPO4 – Polyacrylate capped iron oxide with an YPO_4 shell

PB – Polo-box

PBD – Phosphoprotein-binding domain

Pc – Polo-cap

PHEMA – Poly(2-hydroxyethyl methacrylate)

PhIAT – Phosphoprotein isotope-coded affinity tag

PhIST – Phosphoprotein isotope-coded solid-phase tag

Phos-tag – 1,3-Bis[bis(2-pyridinylmethyl)amino]-2-propanolato dimetal(II) complex

pI – Isoelectric point

PKC – Protein kinase C

PLC-γ1 – Phospholipase C-γ1

Plk – Polo-like kinase

PMF – Phosphopeptide mole fraction

PPIase – Peptidyl-prolyl *cis-trans* isomerase

pSer – Phosphoserine

PTB – Phosphotyrosine-binding

pThr – Phosphothreonine

pTyr – Phosphotyrosine

PyBOP – Benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate

Q_{max} – Maximum binding capacity

R-Smad – Receptor-specific Smad

Rt – Retention time

RT – Room temperature

RT-PCR - Reverse transcription polymerase chain reaction

S/N – Signal-to-Noise

SAM – Self-assembled monolayer

SAX – Strong Anion Exchange

SCX – Strong Cation Exchange

SH2 – Src-homology-2

SIMAC – Sequential elution from IMAC

SW6 – Ser-Gln-Val-Phe-Pro-Trp

SW6-P – pSer-Gln-Val-Phe-Pro-Trp

TCEP – Bond-Breaker tris(2-carboxyethyl)phosphine solution

TEM – Transmission electron microscopy

TEOS – Tetraethyl orthosilicate

TFA – Trifluoroacetic acid

TGF-β – Transforming growth factor-β

THAP – 2′,4′,6′-trihydroxyacetophenone

TIS – Triisopropylsilane

TLC – Thin layer chromatography

TMPTMA - Trimethylol-propane trimethacrylate

TNB^{2-} – 2-Nitro-5-thiobenzoate ion

TW6 – Thr-Gln-Val-Asp-Ala-Trp

TW6-P – pThr-Gln-Val-Asp-Ala-Trp

Washing buffer 1 – 50% ACN/0.5% TFA, 200 mM NaCl

Washing buffer 2 – 50% ACN/0.1% TFA

WD – Trp-Asp dipeptide

YW13 – Tyr-Ala-Gly-Ser-Thr-Asp-Glu-Asn-Thr-Asp-Ser-Glu-Trp

YW13-P – Tyr-Ala-Gly-pSer-pThr-Asp-Glu-Asn-pThr-Asp-Ser-Glu-Trp

YW6 – Tyr-Glu-Glu-Ile-Pro-Trp

YW6-P – pTyr-Glu-Glu-Ile-Pro-Trp

Zn(II)-Dpa – Dinuclear zinc(II)-dipicolylamine

ZrP-pSi - Zirconium phosphonate-modified porous silicon

α-CHCA – α-Cyano-4-hydroxycinnamic acid

γ-Fe_2O_3@xNH_4F.yLuF_3 – Iron oxide core coated with ammonium fluoride and lutetium fluoride

BACKGROUND

The size of the human proteome is at least 10-fold higher than the number of protein-coding genes (≈ 24,000), due to the occurrence of alternative splicing, post-translational modifications, and polymorphisms. Additionally, many proteins exist in different conformational and functional states [1, 2]. Therefore, there is great interest in the development of high-throughput protein-binding detection reagents, which would assist in the specific identification and quantification of numerous disease-related proteins. Affinity proteomics is a rapidly expanding field, and several international initiatives have already been established towards the creation of large collections of well-validated binding reagents. These initiatives include ProteomeBinders (www.proteomebinders.org), Human Protein Atlas (www.proteinatlas.com), HUPO Antibody Initiative (www.hupo.org/research/hai), Antibody Factory (www.antibody-factory.de), among others [2, 3]. The number of affinity binders is continually increasing, and they can be divided in different categories: antibodies; protein scaffolds (e.g. affibodies); peptide ligands (natural or synthetic); nucleic acid aptamers; and small chemical entities (natural or synthetic) [2, 4].

In particular, phosphorylated proteins are involved in complex signalling networks and the establishment of specific protein-protein interactions. Many of these phosphoproteins have unknown functions, and some have been implicated in human diseases. It is fundamental to unveil the interaction landscape of these phosphorylated proteins and their biological binders. In the last decades, a plethora of affinity reagents were developed for the specific binding of phosphorylated proteins and peptides. Their identification by mass spectrometry (MS) analysis poses a particular challenge, due to their low relative abundance in biological samples and to the dynamic character of protein phosphorylation [5]. For the correct identification of phosphorylated species there is the need to enrich their content in biological samples prior to MS analysis.

Enrichment and detection of phosphorylated species can be performed using antibodies which are still the most commonly used affinity molecules in general. However, the phosphate group presents low immunogenicity and is susceptible to cleavage during the immunization process, which hampers the generation of high-quality phosphorylation-state specific antibodies [6]. Most strategies currently used for phosphopeptide enrichment rely on metal coordination of the phosphate moiety, such as immobilized-metal affinity chromatography (IMAC), metal oxide affinity chromatography (MOAC), and even small chemical ligands (e.g. Phos-tag). However, these broad-range methodologies do not provide sufficient diversity.

In view of the interest in studying phosphorylated proteins, and the need to develop tools for the specific enrichment and detection of phosphorylated species, this work focused on the design and production of new specific ligands for the molecular recognition of phosphorylated target peptides.

Two research strategies were followed: the combinatorial synthesis of small chemical ligands and the rational-design of peptide scaffolds (Figure B.1). This constitutes a novel approach in phosphoproteomics with the small and robust affinity ligands presenting potential for higher selectivity when compared with conventional chelating resins.

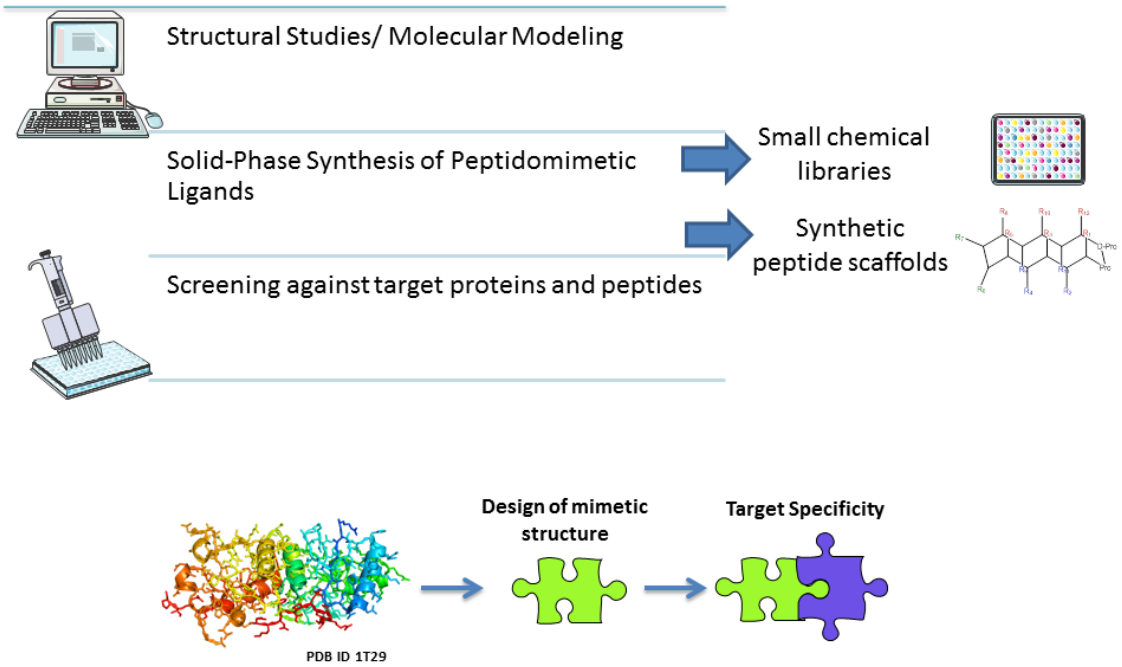


Figure B.1 – Schematic representation of the research strategy followed in the present work.

Bibliography

1. Stoevesandt, O. and M.J. Taussig, *Affinity proteomics: the role of specific binding reagents in human proteome analysis.* Expert Reviews in Proteomics, 2012. **9**(4): p. 401-414.
2. Taussig, M.J., O. Stoevesandt, C.A. Borrebaeck, A.R. Bradbury, D. Cahill, C. Cambillau, A. De Daruvar, S. Dübel, J. Eichler, and R. Frank, *ProteomeBinders: planning a European resource of affinity reagents for analysis of the human proteome.* Nature Methods, 2007. **4**(1): p. 13-17.
3. Stoevesandt, O. and M.J. Taussig, *Affinity reagent resources for human proteome detection: initiatives and perspectives.* Proteomics, 2007. **7**(16): p. 2738-2750.
4. Bourbeillon, J., S. Orchard, I. Benhar, C. Borrebaeck, A. De Daruvar, S. Dübel, R. Frank, F. Gibson, D. Gloriam, and N. Haslam, *Minimum information about a protein affinity reagent (MIAPAR).* Nature Biotechnology, 2010. **28**(7): p. 650-653.
5. Thingholm, T.E., O.N. Jensen, and M.R. Larsen, *Analytical strategies for phosphoproteomics.* Proteomics, 2009. **9**(6): p. 1451-1468.
6. Cai, D., A. Lee, C.-M. Chiang, and T. Kodadek, *Peptoid ligands that bind selectively to phosphoproteins.* Bioorganic & Medicinal Chemistry Letters, 2011. **21**(17): p. 4960-4964.

CHAPTER 1[1]

Platforms for the enrichment of phosphorylated proteins and peptides

Summary

Protein phosphorylation is a complex and highly dynamic process involved in numerous biological events. Abnormal phosphorylation is one of the underlying mechanisms for the development of cancer and metabolic disorders. The identification and absolute quantification of specific phospho-signatures can help elucidate protein functions in signaling pathways and facilitate the development of new and personalized diagnostic and therapeutic tools. This chapter presents a variety of strategies currently utilized for the enrichment of phosphorylated proteins and peptides prior to mass spectrometry (MS) analysis during proteomic studies. The investigation of specific affinity reagents, allied to the integration of different enrichment processes, is triggering the development of more selective, rapid and cost-effective high-throughput automated platforms.

[1] Chapter 1 is partially reproduced from *Batalha, I.L., Lowe, C.R., Roque, A.C.A (2012) "Platforms for enrichment of phosphorylated proteins and peptides in proteomics", Trends in Biotechnology, 30 (2), 100-110.*

Chapter 1 Platforms for enrichment of phosphorylated proteins and peptides in proteomics

1.1. Protein Phosphorylation

The selective and reversible phosphorylation of proteins is a key regulatory mechanism for biological processes, illustrated by the fact that 30-50% of proteins may be phosphorylated at any time [1]. Phosphorylation is recognized as one of the most common post-translational modifications, affecting both the folding and function of proteins. Moreover, phosphorylation events in signaling networks regulate much of the cellular response to external stimuli, and dysregulation in these networks has been linked to multiple disease states (Figure 1.1). There are nine known phosphorylation targets in proteins: Tyr, Ser, Thr, His, Asp, Glu, Lys, Arg, and Cys residues. However, in eukaryotic systems, phosphorylation predominantly takes place at Ser, Thr, or Tyr residues [1, 2].

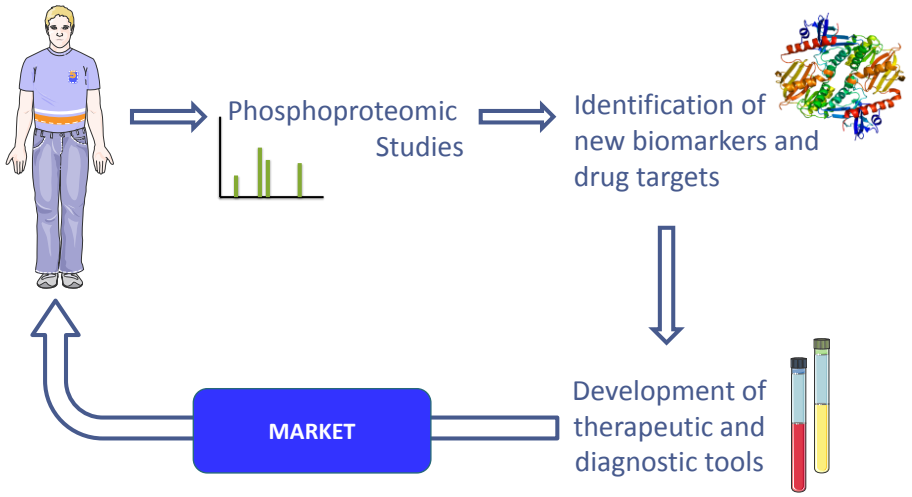


Figure 1.1 – Phosphoproteomics towards the design of personalized medicine and the development of new biomedical tools. Protein phosphorylation networks are implicated in a variety of diseases, including cancer [3], neurodegenerative diseases such as Parkinson's [4] and Alzheimer's [5], and psychiatric disorders such as schizophrenia [6]. Phosphorylated proteins are important biomarkers and drug targets for therapeutic intervention. In fact, the study of real-time signaling activity within cellular specimens (e.g. tumors) helps to profile phosphorylation events to develop personalized therapies and identify patients who might benefit the most from these new therapies. For example, the replacement or combination of nonspecific chemotherapeutics by designed drugs that target specific pathways has proven to be a promising field [7].

1.2. Human phosphoprotein-binding domains (PBDs)

PBDs are involved in the regulation of complex signalling networks by promoting specific protein-protein interactions, assisting the assembly of multi-protein complexes and connecting upstream kinases with their downstream effectors. It is important to bear in mind that, notwithstanding the significance of phosphate recognition in the establishment of protein-protein contacts, some domains also bind their *in vivo* targets in a phospho-independent manner (e.g. 14-3-3, Forkhead-

associated (FHA), etc) [8]. However, this work will be focused exclusively on the interactions of PBDs with their phosphorylated targets. Although most PBDs display extensive structural differences, it has been considered that they may have evolved from common ancesters. As an illustrative example, Mad-homology 2 (MH2) domains have structural similarities with FHA domains and mechanistically resemble Src-homology-2 (SH2) domains [8, 9].

1.2.1. Human phosphoserine (pSer)/threonine(pThr)-binding domains

1.2.1.1. 14-3-3

This was the first pSer/ pThr-binding domain identified and is expressed in all eukaryotic cells. 14-3-3 proteins are associated with a variety of cellular processes, such as cell cycle progression, DNA damage responses, metabolism, receptor recycling and trafficking, protein degradation and apoptosis, among others, although their mechanistic action is not yet fully understood [10, 11]. They appear to regulate the catalytic activity of some ligands (e.g. Raf), the interaction between their bound ligands and other molecules within the cell (e.g. pro-apoptotic Bcl-2 family protein BAD), and controlling the subcellular localization of bound proteins (e.g. Cdc25) [11]. They present three fundamental modes of action: (i) conformational change of the ligand; (ii) physical obstruction of a specific region on the target ligand, blocking protein-protein and protein-DNA interaction sites, and protecting their target ligands against dephosphorylation and/or degradation; (iii) anchoring of two target proteins in the vicinity of one another, promoting protein-protein interactions [10, 12].

In terms of structure, 14-3-3 proteins are C2-symmetric dimers (either homo- or heterodimers) [10, 13]. Each monomeric subunit is composed of nine antiparallel α-helices, connected by short loops. The four N-terminal helices (A-D) interact by salt bridges, hydrogen bonds, and van der Waals interactions with the opposite subunit, forming a central 35 Å wide and 20 Å deep channel, where the target ligand binds (Figure 1.2A) [13, 14]. Binding occurs in an amphipathic groove along each edge of the channel. A basic pocket, consisting of Lys49, Arg56, and Arg127, along with Tyr128 (from helices C and E), are responsible for phosphate binding, through the formation of five ionic bonds and a hydrogen bond to three of the phosphate oxygen atoms [10, 12, 13, 15]. The phosphopeptide main chain is positioned in an extended conformation by hydrophobic interactions with Leu172, Val176, Leu216, Ile217, Leu220, Leu227, and Trp228 from the C-terminal helices, as well as hydrogen bonding between the peptide NH and CO backbone groups and side chains of residues in helices E, G and I, with Asn173 and Asn224 playing a crucial role (Figure 1.2B) [10, 13-15]. The presence of a

proline at the C-terminal two residues to the phosphorylated residue redirects the peptide chain away from the binding groove [10, 13].

Two consensus binding motifs were found by oriented peptide library screening: Arg-(Ser/Ar)-(+)-pSer-X-Pro and Arg-X-(Ar/Ser)-(+)-pSer-X-Pro, where Ar is an aromatic residue, + is a basic residue, and X is any residue (commonly Leu, Glu, Ala and Met) [10, 14, 16].

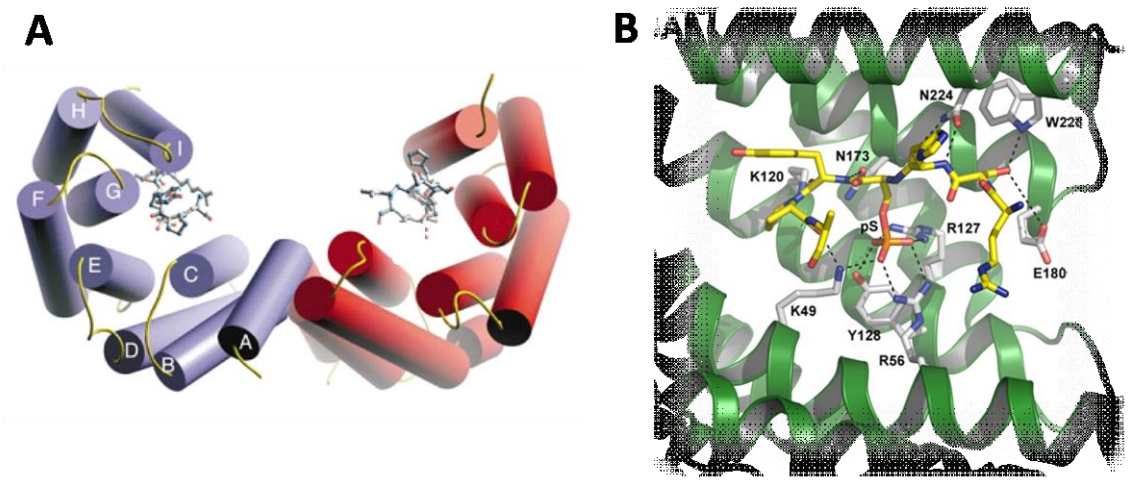


Figure 1.2 – Schematic representation of 14-3-3 domains binding to a consensus peptide motif. A) 14-3-3 proteins form cup-shaped dimeric structures. Each monomer (in red and blue) contains nine α-helices (A-I) linked by short loops. In this representation, two phosphorylated peptides are simultaneously binding to each monomer, but a unique target protein with multiple phosphorylation sites could also bind in an identical manner. B) The phosphate group establishes five ionic bonds and one hydrogen bond with Lys49, Arg56 and Arg127 and Tyr128. Peptide backbone interacts with Asn173 and Ans224 through hydrogen bonding. Nitrogen atoms are colored in blue, oxygen atoms in red, and phosphate group in orange. (Images A and B are reproduced from references [13] and [10], respectively.)

1.2.1.2. Brca1 C-terminal (BRCT)

BRCT domains consist of 85 to 100 amino acids and have been observed as single, tandem or multiple repeats in a variety of proteins, namely BRCA1 (where it was originally identified), 53BP1, XRCC1, DNA ligases III and IV, among others [17, 18]. A single protein can contain up to eight of these domains, which are of pivotal importance in cellular DNA damage response, being involved in cell cycle control and checkpoint-mediated repair. Mutations in BRCA1 sequences encoding the BRCT domains have been highly associated with breast and ovarian cancers, by inhibiting phosphopeptide recognition [19-21].

Structurally speaking, each BRCT domain comprises a four-stranded parallel β-sheet flanked by three α-helices (α1 and α3 on one side and α2 on the other), presenting highly conserved structures and selectivity for the sequence pSer-X-X-Phe, where X is any amino acid (Figure 1.3A). Although these

domains recognize both pSer and pThr, the first one binds with much higher affinity [21]. Most conserved residues are located in the central β-sheet, in helix α3 and at the C-terminus of helix α1, while helix α2 is poorly conserved [19]. Phosphopeptide recognition usually requires the existence of two repeats, with the peptide being positioned in a groove at the center of the dimer. The inter-BRCT linker is a variable region, both in sequence and length, and it is not clear if its function is solely to properly position each domain or if it is also implicated in the establishment of protein-protein interactions [17, 19]. pSer interacts with the N-terminal BRCT repeat, whereas Phe is recognized by a hydrophobic pocket at the N- and C-terminal repeats interface. Both the main chain NH of Gly1656 and the OH of Ser1655 of the BRCA1-BRCT domain interact by hydrogen bonding with the phosphate group, with the OH group of Ser1655 being further stabilized by hydrogen-bonding with Thr1700. The phosphate group also establishes salt bridges with Lys1702, which in turn interacts by hydrogen bonding with the main chain CO of Val1654 and Asn1678 (Figure 1.3B) [19, 22]. This type of coordination of phosphate oxygen atoms by main chain NH groups, Ser OH groups and positively charged side chains resembles the binding mode of 14-3-3 and FHA domains [19]. The Phe-binding pocked consists of Leu1701, Phe1704, Met1775 and Leu1839 at the bottom, while Arg1699, Asn1774 and Arg1835 are positioned at the sides. Particularly, Arg1699 is of extreme importance for the positioning of Phe, forming hydrogen bonds through its main chain CO with the main chain NH of Phe, while its guanidinium side chain interacts though salt bridges with the main chain CO of Phe [19, 23]. Moreover, mutations in Met1775 destroy the ability of BRCT-containing proteins to interact with their targets, and are often observed in patients with cancer [19, 20].

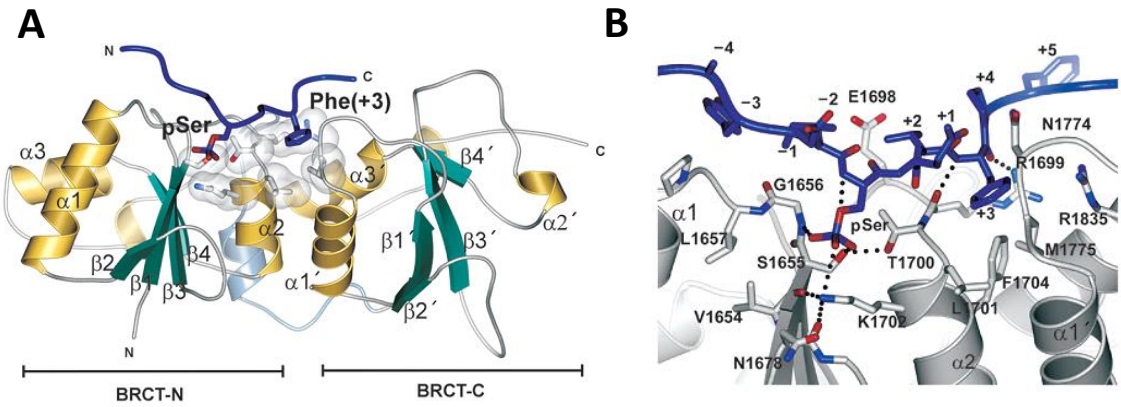


Figure 1.3 – BRCT domain binding to pSer-X-X-Phe peptide motif. A) Each BRCT domain is composed of four β-sheets and three α-helices. Phosphopeptide recognition typically requires the presence of two domains, with the phosphorylated residue being positioned in the N-terminal BRCT, while the Phe recognition pocket is located in the interface between the domains. B) Detailed view of phosphopeptide binding to BRCT domains. pSer-binding pocked consists of three critical amino acids: Ser1655, Gly1656 and Lys1702. Phe in position (+3) is mainly stabilized by non-polar interactions with Met1775 and backbone interactions with Arg 1699. (Reproduced from reference [19].)

1.2.1.3. FHA

FHA domains are pThr-binding proteins composed of approximately 140 amino acids [11, 24]. They are found in many proteins, including kinases, phosphatases, transcription factors, metabolic enzymes and kinesin-like motors [8, 25]. In humans, they have been associated with cancer and particularly with a variant form of cancer-prone Li-Fraumeni syndrome [26].

FHA domains are organized in a 10- or 11- stranded β-sandwich, in a way that closely resembles MH2 domains (Figure 1.4A) [8, 24]. The loops connecting the strands are essential for pThr binding, which occurs at only one end of the domain [24, 27]. Phosphate recognition is made by two conserved residues, an Arg of β3-β4 loop and a Ser of β4-β5 loop, and an Arg, Lys or Asn of the β4-β5 loop, through hydrogen-bonding and salt-bridging (Figure 1.4B). The γ-methyl group of pThr establishes structurally conserved van der Waals interactions, which explains the discrimination of pThr over pSer [8, 25]. However, the phosphorylated residue is not the only one conferring specificity. FHA domains bind in a sequence-specific manner from pThr-4 to pThr+3 positions, with the highest selection in the pThr+3, which is preferably an Asp, a Leu or Ile, depending on the protein. Residues from loops β4-β5, β6-β7 and β10-β11 are mostly involved in this interaction [8, 24].

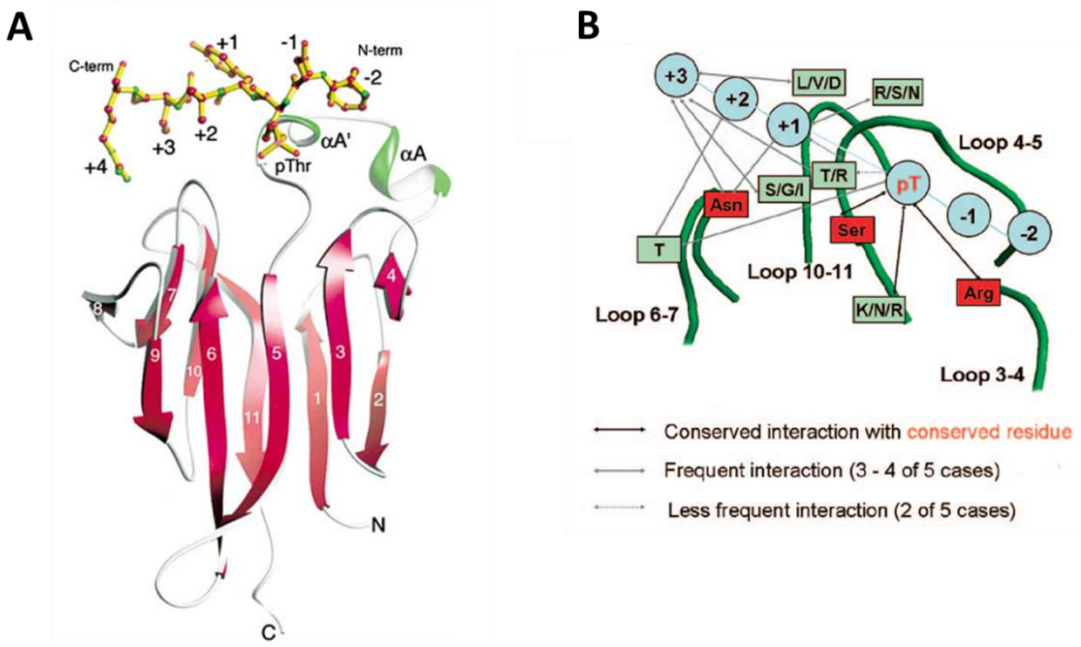


Figure 1.4 – Structural details of a pThr-containing peptide binding to FHA domains. A) FHA Domains consist of 10- or 11-stranded β-sandwich (colored in pink). pThr-containing motifs interact with FHA domains via β-strand-connecting loops (colored in white). B) pThr interacts by hydrogen-bonding and salt-bridging with a conserved Arg of β3-β4 loop, a conserved Ser of β4-β5 loop, and an Arg, Lys or Asn of the β4-β5 loop. A conserved Asn of β6-β7 loop contributes for the specific recognition of residues at pThr+1 and pThr+3 positions (Images A and B were reproduced from references [8] and [25], respectively).

1.2.1.4. MH2

The transforming growth factor-β (TGF-β) signalling pathway regulates numerous cellular processes, both during embryogenesis and in adult life, and the emergence of some diseases such as cancer, atherosclerosis, renal and cardiac fibrosis, Marfan syndrome, among others, has been associated with the disruption of this pathway [28, 29]. TGF-β signalling pathway is mediated mainly by transmembrane receptors with Ser/Thr and weaker Tyr kinase activities, and also cytoplasmic effector proteins named Smad [29, 30]. Receptor-specific Smads (R-Smads) and the common mediator Smad - Smad4 - are composed of two highly conserved domains, the Mad-homology 1 (MH1) domain at their N-terminus and the MH2 domain at their C-terminus, which are connected by a proline-rich variable linker. The MH1 domain is mainly a DNA-binding domain, although it also binds to regulatory proteins and is thought to be a negative regulator of the MH2 domain. On the other hand, MH2 domains mediate interactions with a wide range of proteins, including receptors, other Smad proteins, nuclear transcriptional co-activators and co-repressors [28, 29]. MH2 domains possess a large hydrophobic surface, which facilitates their interaction with proteins that are unrelated both in terms of sequence and structure [28].

In simplified terms, the receptor kinases suffer oligomerization upon ligand (e.g. TGF-β) binding and phosphorylate a R-Smad protein (e.g. Smad1, -2, -3, -5) at its C-terminal MH2 domain in a highly conserved region, Ser-Ser-X-Ser, where X is any amino acid. Phosphorylation induces first homo-trimerization of R-smads and then hetero-oligomerization between R-Smad and Smad4, which facilitates R-Smad-Smad4 complex to enter the nucleus where it binds to certain promoters and regulates the transcription of a variety of genes (Figure 1.5A) [9, 31].

MH2 domains consist of approximately 200 amino acids and a central β-sandwich surrounded by a three-helical bundle at one side and a loop/helix region at the other (Figure 1.5B) [29, 30]. Phosphorylation of R-Smads occurs in two of the three Ser residues of the consensus sequence (e.g. Ser465 and Ser467 in Smad2). The negatively charged C-terminus of R-Smad interacts by hydrogen-bonding and van der Waals forces with a positively charged loop-strand pocket of either R-Smads (in homo-oligomerization) or Smad4 (in hetero-oligomerization), which is formed by the L3 loop and the β-strand B8. Four conserved amino acids, both in R-Smads and Smad4, play a prominent role in phosphate coordination: Lys375 of β-strand B8 and Lys420, Tyr426 and Arg428 of L3 loop (Figure 1.5C). Tyr426 and Lys375 make three hydrogen bonds to the phosphate group of Ser465, whereas Lys420 and Arg428 hydrogen-bond Ser467. Ser467 is further stabilized by hydrogen-bonding between its carboxylate group and Lys420, Arg427 and Arg428 side chains. Intramolecular contacts

occur between Ser465 phosphate group and the backbone amide group of Met466, and Ser467 backbone amide and phosphate group with Ser464 side chain. The interaction between Smads is further strengthened by van der Waals contacts [31]. Arg378, which is present in Smad4 but not in R-Smads, contributes for the preferential hetero-oligomerization of Smads and also for the coordination of phosphorylated Ser467 by hydrogen-bonding [32].

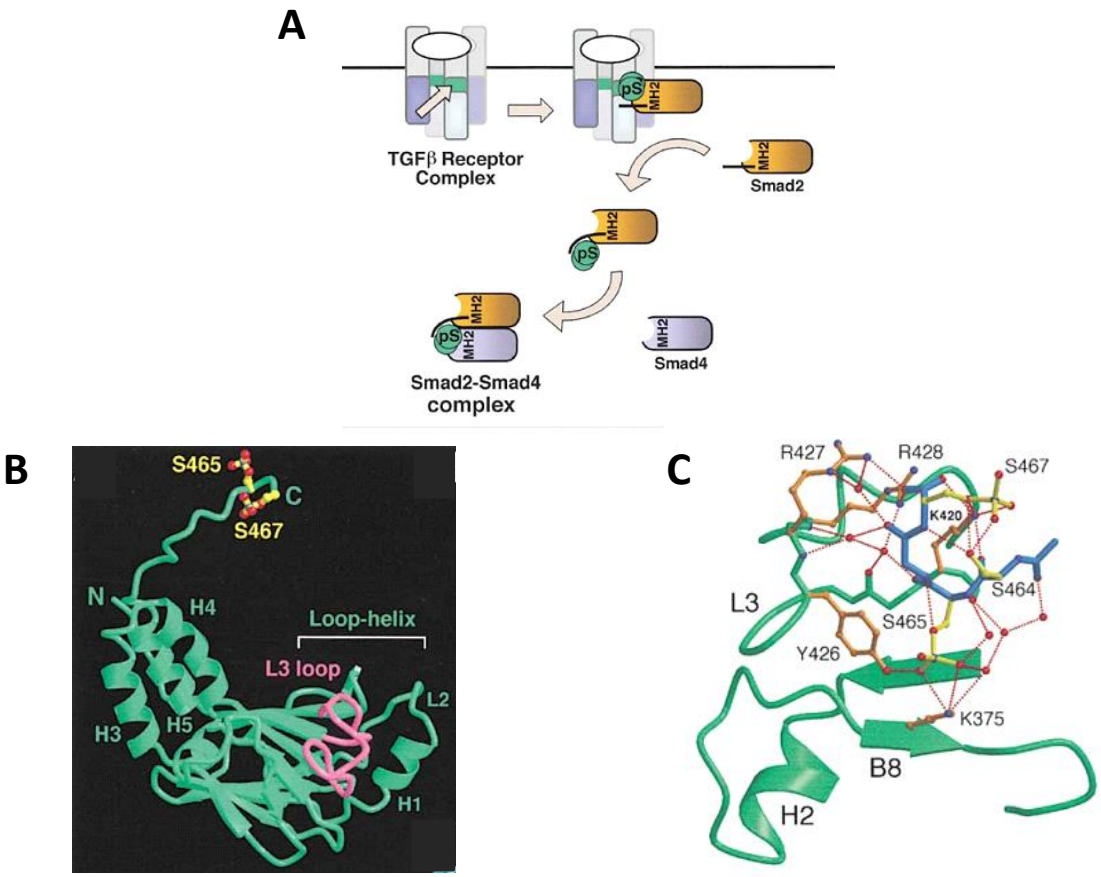


Figure 1.5 – Mechanistic and structural insights of Smad2 MH2 domain. A) Diagram of TGF-β signalling pathway. TGF-β binds to a transmembrane receptor complex inducing phosphorylation of Smad2 MH2 domain, which in turn binds to Smad4 MH2 domain. Smad2-Smad4 complex then translocates to the nucleus where it regulates gene expression. B) MH2 domains consist of a central β-sandwich bordered by three helices on one side and three loops and one helix on the other side. C) Phosphate groups of Ser465 and Ser467 are coordinated by four conserved residues, Lys375 on β-strand B8 and Lys420, Tyr426 and Arg428 on L3 loop. (Reproduced from reference [31].)

1.2.1.5. Polo-box (PB)

Polo-like kinases (Plks) are a family of Ser/Thr protein kinases involved in all stages of mitosis, cytokinesis, and cell cycle checkpoint response to genotoxic stress [33, 34]. Humans hold five Plk family members: Plk1, Plk2, Plk3, Plk4 and Plk5. Plk1 to Plk4 are all involved in cell cycle progression, while Plk5 is exclusively implicated in non-proliferative processes, such as neuronal differentiation [35].

Plks have an N-terminal kinase domain (absent in human Plk5) and a C-terminal non-catalytic region – the PB domain, which is fundamental for Plk localization and function. This domain comprises two Polo-boxes in Plk1, Plk2, Plk3 and Plk5, and a single Polo-box in Plk4. PB domains of Plk1 and Plk2 recognize a specific phosphorylated sequence, Ser-(pSer/pThr)-(Pro/X), where X is any amino acid. In contrast, the PB domain of Plk3 does not act exclusively in a phospho-dependent manner and in fact presents low affinity for phosphopeptides. PB domains of Plk4 and Plk5 work completely as phospho-independent modules [11, 33, 36].

The PB domain is a 12-stranded β-sandwich flanked by three α-helices. It is composed of two Polo-boxes (PB1 and PB2) and a linker between the kinase domain and PB1 (Polo-cap or Pc). Each Polo-box consists of 70-80 amino acids forming a 6-stranded anti-parallel β-sheet flanked by one α-helix. PB1 and PB2 are linked together through a conserved linker sequence (L2). The Pc region comprises one α-helix (αA), a loop and a short 3_{10} helical structure that is attached to β1 strand of PB1 by a short linker (L1). The Pc folds around PB2 tethering it to PB1.

Phosphopeptide binding occurs at one edge of the interface between PB1 and PB2 in a conserved positively charged cleft (Figure 1.6A) [33, 34]. Binding of phosphate moiety involves not only protein-peptide interactions, but also extensive hydrogen-bonding to an ordered lattice of water molecules. The only residues that directly interact with the phosphate moiety are one conserved His and one conserved Lys from PB2 (Figure 1.6B) [11, 33, 34]. Plk1 structure indicates that there is a highly conserved Ser residue at pThr-1 position, which hydrogen-bonds to the main chain of conserved Trp414 and to the main chain carbonyl of conserved Leu491 via a water molecule. Moreover, it establishes significant van der Waals interactions with Trp414 indole group (Figure 1.6C). The preference of a Pro residue at pThr+1 position has been attributed to the fact that it redirects the peptide backbone, facilitating the positioning of the pThr+2 residue near the binding surface. There is some debate in literature as to whether there is a preference for a pThr over a pSer residue in the consensus motif [33, 34].

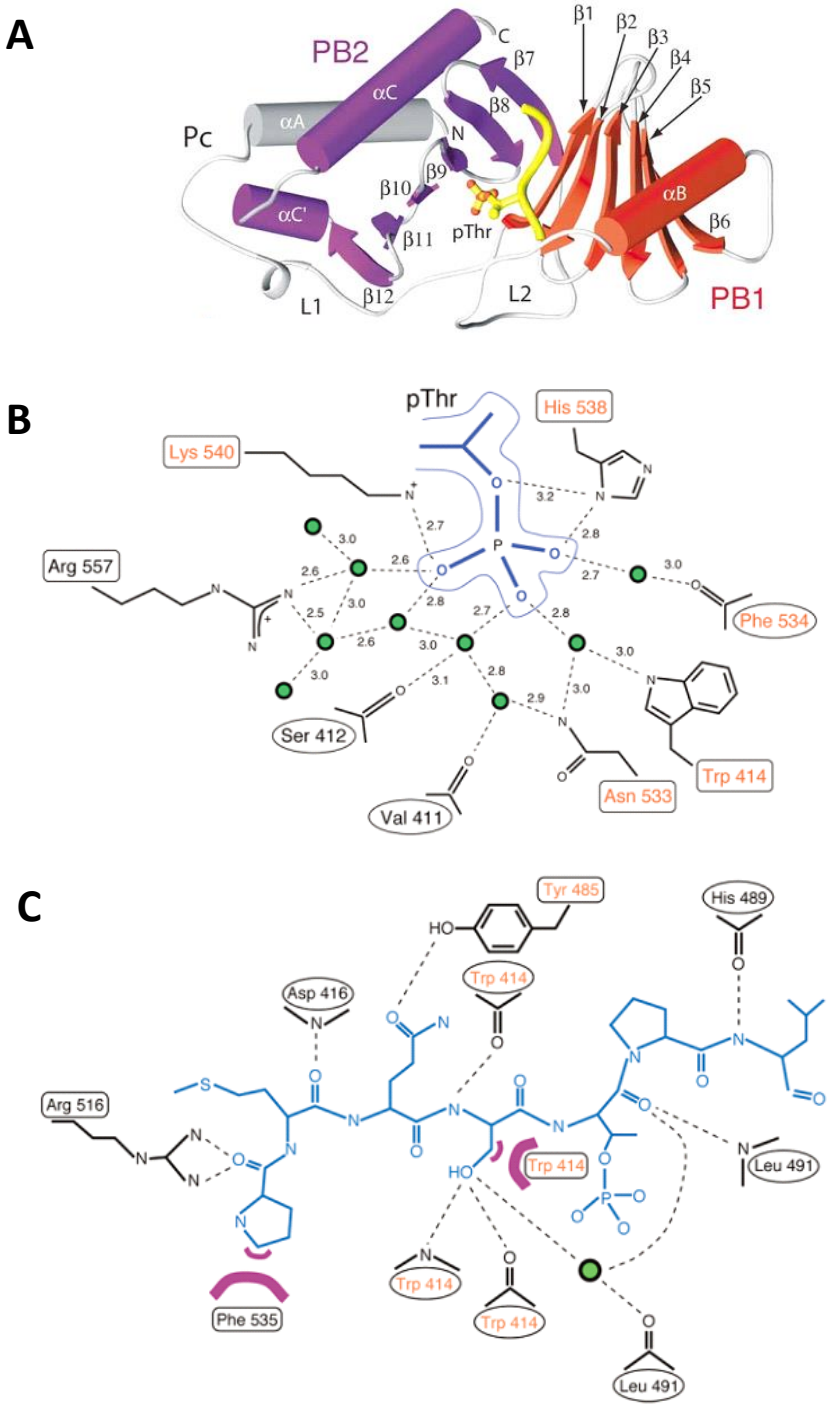


Figure 1.6 – PB domain/ phosphopeptide interactions. A) PB domains are composed of two Polo-boxes (PB1 and PB2), each of them consisting of six β-strands and one α-helix, and an N-terminal region (Pc) which surrounds PB2. The Pc region is connected to PB1 by one loop (L1) and the polo-boxes are connected to each other through another loop (L2). Phosphopeptide binding occurs at the interface between PB1 and PB2. PB1 is coloured red, PB2 is coloured purple and Pc is coloured grey. B) Phosphate moiety is coordinated by a well-defined network of water molecules. The unique residues that make direct contacts to phosphate are His538 and Lys540. C) Interactions between the phosphopeptide (coloured blue) and the Polo-box domain Ser at position pThr-1 establishes highly conserved interactions with Trp414 and Leu491. Hydrogen bonds are represented by dotted lines; van der Waals contacts by purple crescents, and water molecules by green circles. (Reproduced from reference [33].)

1.2.1.6. WD-40

WD-40-containing proteins are found in all eukaryotes and some prokaryotes, and are known to be involved in a wide variety of cellular processes, such as signal transduction, cell cycle control, RNA synthesis and processing, chromatin assembly, cytoskeletal assembly, regulation of vesicular formation and vesicular trafficking, programmed cell death, etc. [37, 38]. Despite this diversity, there is one shared function among WD-40 proteins: all of them support the assembly of multi-protein complexes by acting as protein scaffolds, allowing the simultaneous interaction with several proteins [37]. The WD repeat consists of 44 to 60 amino acids, generally containing a Gly-His (GH) dipeptide 11 to 24 amino acids from its N-terminus and a Try-Asp (WD) dipeptide at its C-terminus, although none of the dipeptides are completely conserved. There is a conserved 40 amino acid core sequence between GH and WD, which is why it is designated as WD-40 repeat. WD-40 proteins possess a minimum of 4 WD-40 repeats and up to 16, folding into a β-propeller structure. Each WD-40 repeat consists of a 4-stranded antiparallel β-sheet. The propeller blades are composed of 3 β-strands of one repeat and one β-strand of the neighbouring repeat, which provides structural stability (Figure 1.7A) [37, 38].

WD-40 domains are able to recognize specific phosphorylated substrates: the WD-40 domain of Cdc4 binds to the (Leu/Ile)-(Leu/Ile/Pro)-pThr-Pro-X motif, where X is any amino acid except Arg and Lys [11, 39]; and the WD-40 domain of β-TrCP binds to the consensus sequence Asp-pSer-Gly-X-X-pSer in IκBα, β-catenin and Vpu, and to Asp-pSer-Gly-X-X-X-pSer in ATF4, where X is any amino acid [40, 41].

In the WD-40 domain of Cdc4, the phosphate moiety of the substrate is coordinated by hydrogen-bonding to the side chain of Tyr548, and by electrostatic interactions with the guanidinium group of three Arg residues (Arg485, Arg467 and Arg534). The conserved Pro residue at pThr+1 position is positioned in a hydrophobic pocket composed of Trp426, Thr441, and Thr465 (Figure 1.7B) [39].

The WD-40 domain of β-TrCP binds to the first phosphorylated residue of β-catenin (pSer33) and Vpu (pSer52) in a similar way: the phosphate moiety hydrogen-bonds to the side chain OH groups of Tyr271, Ser309, and Ser325, and interact electrostatically with the guanidinium group of Arg285. The second phosphorylated residue - pSer37 in β-catenin and pSer56 in Vpu – establishes fewer interactions, and with different amino acids. pSer37 phosphate group establishes electrostatic interactions with the side chain of Arg431, and hydrogen-bonds to the side chain of Ser448 and the backbone amide of Gly432. pSer56 is located within a positively charged pocket composed of Arg367, Arg390, Arg431 and Lys365. The conserved Asp residue at pSer-1 position of β-catenin makes a hydrogen-bond with Arg474 and Tyr488, while the Asp of Vpu hydrogen-bonds to Arg474, Arg521,

and Tyr271. Gly at pSer+1 position packs with two Leu residues of β-TrCP both in β-catenin and Vpu (Figure 1.7C) [40, 41].

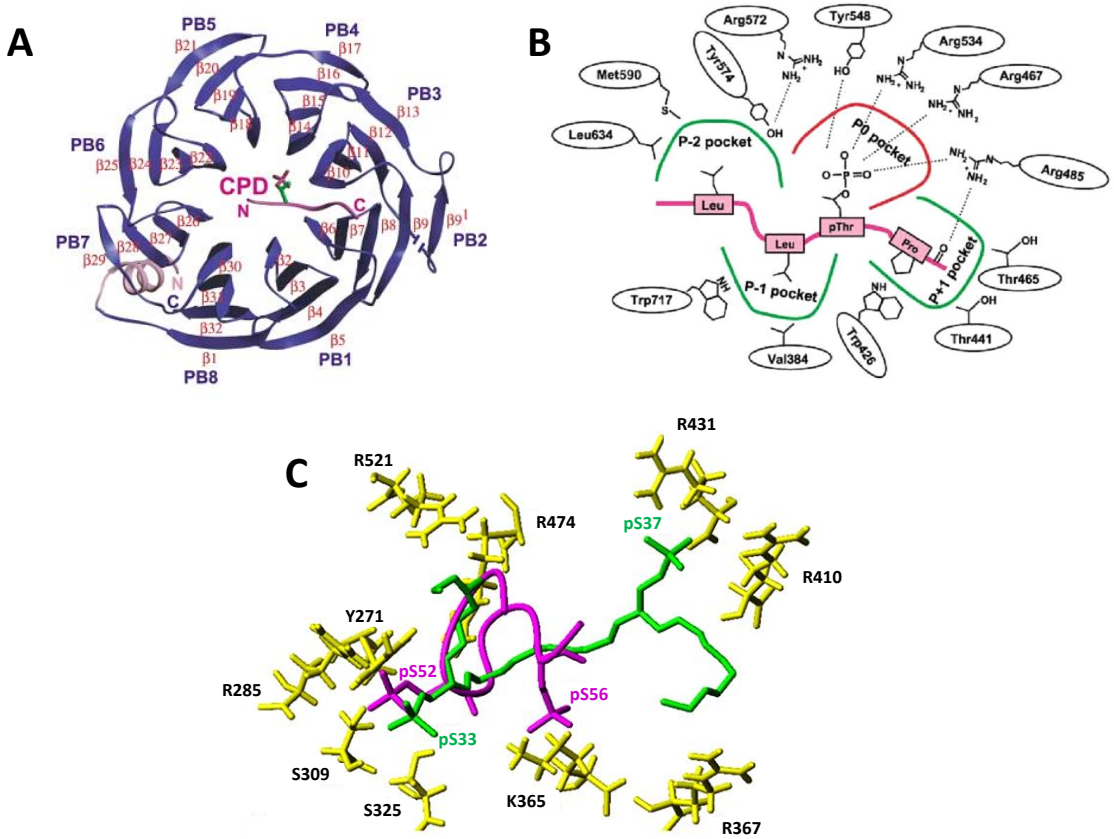


Figure 1.7 – WD-40 domains: Structure and recognition of phosphorylated substrates. A) β-propeller structure of WD-40 domain of Cdc4 binding to Cdc4-phosphodegron (CPD). Cdc4 is composed of eight propeller blades (PB1 to PB8), where each blade consists of three β-strands of one WD-40-repeating unit and another β-strand of the adjacent repeat. B) Interaction sites between Cdc4 WD-40 domain and CPD. The phosphate binding pocket is mainly composed by three Arg residues, which establish electrostatic interactions with the phosphate, and a Tyr which hydrogen-bonds to the phosphate through its side chain. Pro at pThr+1 position and the hydrophobic residues at positions pThr-1 and pThr-2 are all mainly sited within a hydrophobic cluster. Pro establishes an additional contact with Arg485, through van der Waals interactions with its side chain and hydrogen-bonding to its main chain carbonyl. C) Binding of β-TrCP to β-catenin (in green) and Vpu (in pink). The first pSer coordinates in a similar way in both substrates, establishing conserved interactions with Arg285, Tyr271, Ser309, and Ser325. The second phosphorylated residue sits near one Arg in β-catenin and one Lys in Vpu. (Images A and B were reproduced from reference [39], and image C was reproduced from [41].)

1.2.1.7. WW

WW domains are 38 to 40 amino acid modules widely spread in a variety of signalling proteins and that are known to bind proline-rich sequences. They play a multiplicity of roles in cell cycle control, transcription, RNA processing, ubiquitin ligation, protein trafficking, control of the cytoskeleton, receptor signalling, etc., and have been associated with diseases such as cancer, Liddle's syndrome, Rett syndrome, Duchenne and Becker muscular dystrophies, Huntington's disease, and Alzheimer's disease [42, 43].

They can be divided into five classes: Group I, Group II, Group III, Group IV and Group V, which bind to Pro-Pro-X-Tyr (where X is any amino acid), Pro-Pro-Leu-Pro, Pro-Gly-Met, (pSer/pThr)-Pro, and Pro-Arg, respectively [44]. However, this classification has been rather controversial [42]. Here we are going to focus on Group IV, since it is the one recognizing phosphorylated substrates. The crystal structure of the WW-containing protein Pin1 in complex with the doubly phosphorylated Tyr-pSer-Pro-Thr-pSer-Pro-Ser peptide of the carboxy-terminal domain of RNA polymerase II has been solved and will be used here as an example [44]. WW domain adopts a three-stranded antiparallel β-sheet structure and is located at the N-terminus of Pin1, which has a peptidyl-prolyl *cis-trans* isomerase (PPIase) domain at its C-terminus (Figure 1.8A). The PPIase domain catalyse the (pSer/pThr)-Pro bond isomerization, while WW domains are only able to bind the *trans* isomer of their phosphorylated substrates. The crystal structure of the Pin1-phosphopeptide complex reveals that the WW domain establishes interactions with the second phosphorylated Ser residue through a Ser and an Arg in the β1/β2 loop and a Tyr in the β2 strand. Two aromatic residues, a Tyr and a Trp, establish additional van der Waals interactions with Thr4, pSer5 and Pro6 of the peptide (Figure 1.8B) [13, 44].

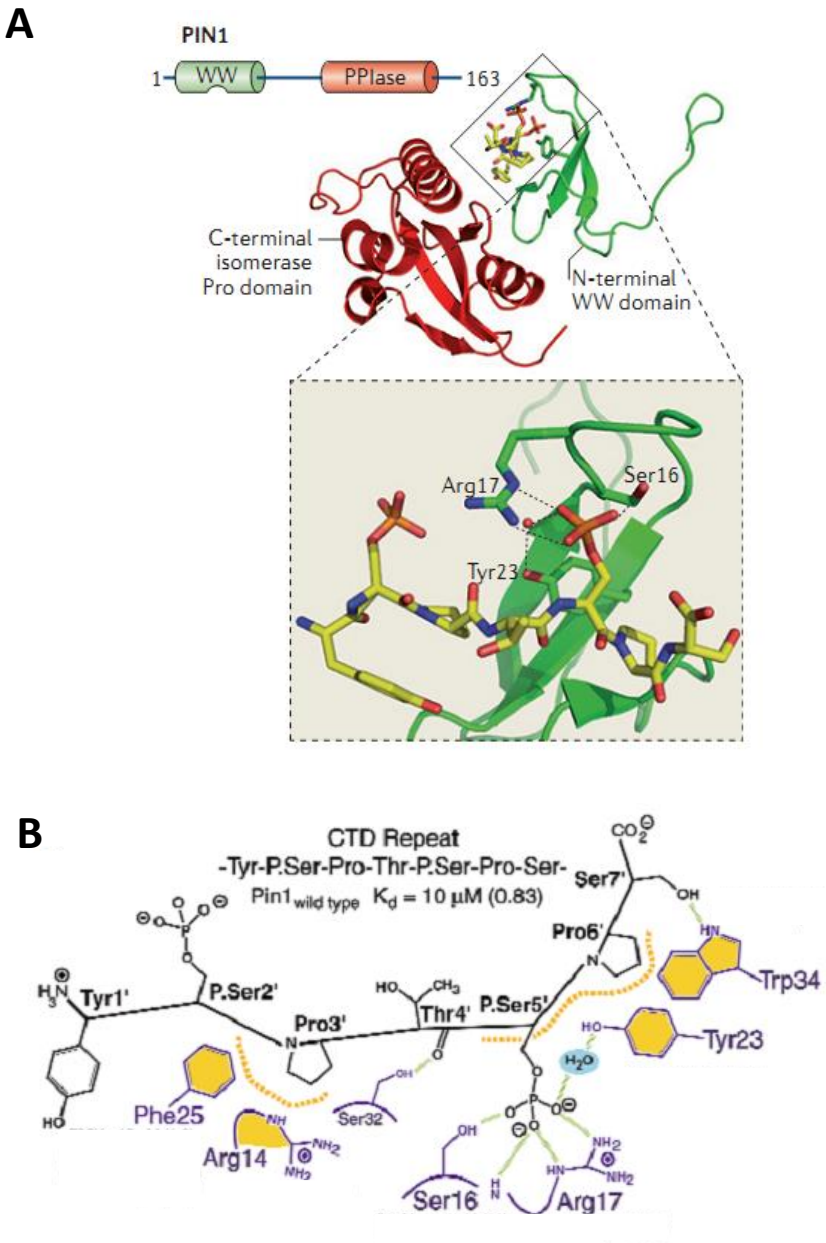


Figure 1.8 – Pin1 WW domain complexed with Tyr-pSer-Pro-Thr-pSer-Pro-Ser peptide of the carboxy-terminal domain of RNA polymerase II (CTD peptide). A) Pin1 is composed by a three-stranded antiparallel β-sheet structured N-terminal WW domain and a C-terminal PPIase, connected by a linker. B) The CTD peptide binds to WW domain through its second phosphorylated Ser residue (pSer5'). The phosphate group binds to the side chains of Ser16 and Arg17 though hydrogen-bonding. Arg17 makes an additional bond to the phosphate through its backbone NH group. The phosphate group interacts with the side chain of Tyr23 by mediation of a water molecule. Both Tyr23 and Trp34 establish van der Waals contacts with the backbone of Thr4' and pSer5' and the ring atoms of Pro6' of the CTD peptide. Hydrogen bonds are represented by green dashed lines and van der Waals surfaces by gold dotted curves. Amino acids of the CTD peptide are labelled black and with an apostrophe for clarity. (Images A and B were reproduced from references [24] and [44], respectively.)

1.2.2. Human phosphotyrosine (pTyr)-binding domains

1.2.2.1. SH2

SH2 domains consist of around 100 amino acids, being the first domains recognized to behave as phosphorylation-dependent binding modules. Among SH2-containing proteins are kinases, phosphatases, transcription factors, adaptor proteins, cytoplasmic signalling molecules, among others [45]. Tyr phosphorylation is a general recognition element for SH2 domains (although a few have been reported to bind in a phospho-independent mode), but they discriminate among their substrates depending on which amino acid residues are positioned carboxy-terminal to pTyr. This allows particular SH2 proteins to be recruited and trigger specific signalling pathways [45, 46]. In fact, SH2 domains can be divided in three main groups, depending on substrate specificity: Src-like, phospholipase C-γ1 (PLC-γ1)-like, and Grb2-like, which select for pTyr-Glu-Glu-Ile, pTyr-hydrophobic-X-hydrophobic (where X is any amino acid), and pTyr-X-Asn-X (where X is commonly a Val residue) motifs, respectively [47-49].

SH2 domains fold into a β-meander containing one four-stranded and one three-stranded β-sheet, which are connected to each other by one β-strand. The four-stranded β-sheet is capped by two additional α-helices. The phosphopeptide binds in an extended conformation perpendicular to the four-stranded β-sheet. The pTyr residue binds at one side of the β-sheet, with two Arg residues (one from αA and one from βB) playing a central role in phosphate coordination. The Arg from αA together with side chain and backbone atoms from βD strand residues further stabilize the aromatic ring of pTyr. Residues from pTyr+1 to pTyr+5 are positioned in the opposite side of the β-sheet, binding to atoms from the loops between αB and βG (loop BG) and between βE and βF (loop EF) (Figure 1.9). These last interactions depend on the class of SH2 domains [45, 50].

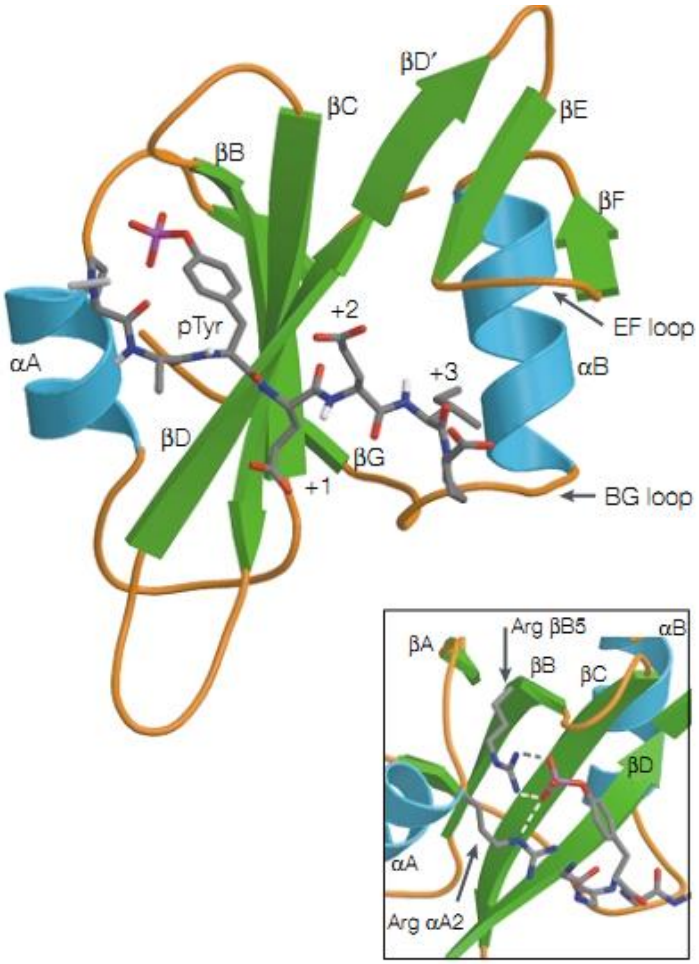


Figure 1.9 – Structure and binding specificity of SH2 domains. SH2 domains fold into a β-meander structure, composed of four stranded and three-stranded β-sheets, which are connected by one β-strand, the N-terminus of which is denominated βD and the C-terminus of which is denominated βD'. Two α-helices (αA and αB) flank the four-stranded β-sheet. The phosphopeptide sits perpendicular to this β-sheet, with the pTyr interacting by hydrogen-bonds with two Arg residues (Arg αA2 and Arg βB5) at one side of the β-sheet. The aromatic group of pTyr is stabilized by an amino-aromatic bond to Arg αA2 and other interactions with residues from βD strand. The residues from pTyr+1 to pTyr+5 are located on the opposite side on the β-sheet, and make contacts to loops BG and EF. (Reproduced from reference [45].)

1.2.2.2. Phosphotyrosine-binding (PTB)

PTB domains may be found in a wide range of adaptor and scaffold proteins, being implicated in cell growth, immune response, neural development, tissue homeostasis, etc. Mutations on PTB domains have been linked to Alzheimer's, diabetes, and cardiovascular diseases [51]. All PTB domains share the same structural conformation, folding into a 7-stranded β-sandwich which is flanked on its C-terminus by one α-helix, although they have non-conserved primary sequences (Figure 1.10A) [45, 51, 52]. PTB domains have been classified in three main classes based on their structural and functional features: Shc-like, IRS-like and Dab-like. The canonical recognition motif is Asn-Pro-X-

(Tyr/pTyr). Both Shc-like and IRS-like domains recognition is dependent on the presence of a phosphorylated Tyr residue on the consensus sequence, while Dab-like domains recognize the non-phosphorylated sequence (or even a Phe instead of a Tyr). In the three cases, the peptide binding pocket is composed of residues from β5 strand and the C-terminal α-helix, but the recognition is made by different amino acids. In IRS-like domains the peptide binding pocked consists of two Arg residues (Figure 1.10B). For Shc-like domains, the phosphate group is coordinated by hydrogen-bonding to two Arg residues and one Lys (Figure 1.10C). Dab-like domains bind to their target sequences driven by the establishment of hydrogen-bonds and hydrophobic interactions with amino acids on β5 strand, rather than recognition of the Tyr residue itself (Figure 1.10D) [51]. Other PTB domains recognizing peptide sequences different from the canonical Asn-Pro-X-(Tyr/pTyr) motif have been reported, and most PTB domains are known to bind phospholipids as well, enlightening the fact that there are no universal patterns of recognition [51, 53].

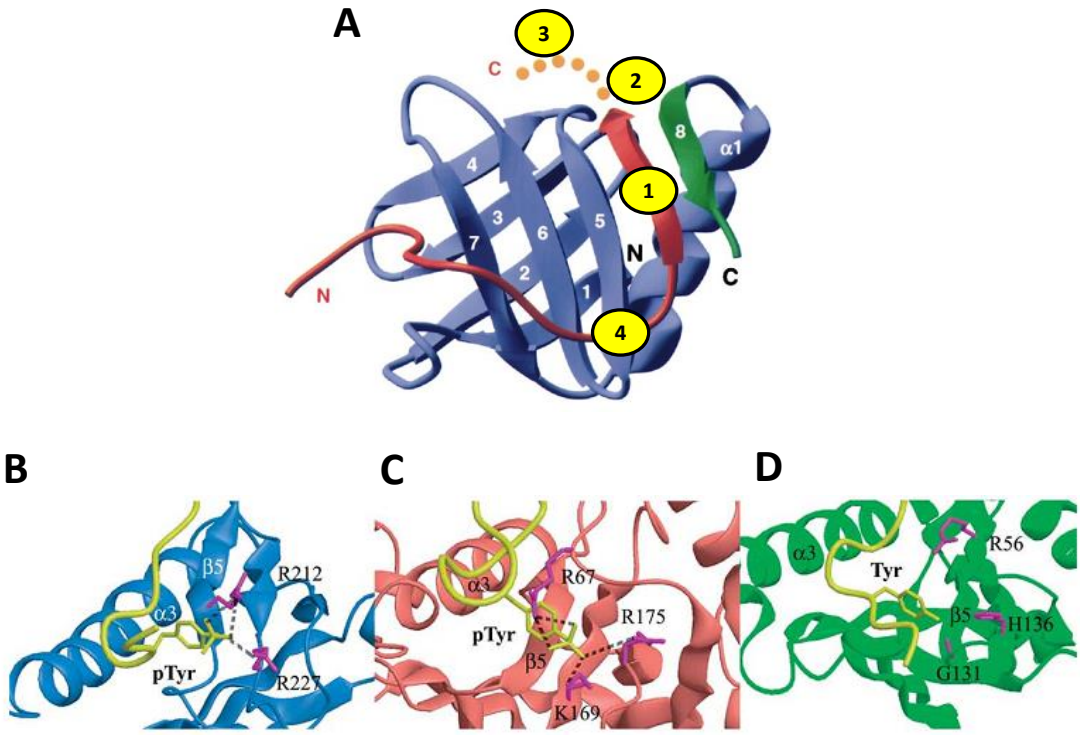


Figure 1.10 – Structure and binding-specificity of PTB domains. A) PTB domains are seven-stranded β-sandwiches flanked by one C-terminal α-helix (α1). A1 – The peptide folds into a β-strand conformation that establishes antiparallel interactions with the β5 strand of PTB. A2 – The peptide forms a type-I β-turn at its consensus motif. A3 – Tyr/pTyr-binding pocket. A4 – N-terminal residues interact by hydrophobic contacts with β5-strand and C-terminal α-helix. B) Phosphate coordination is achieved through contacts with two Arg residues in IRS-like domains. C) Two Arg residues and one Lys form a triangular network of hydrogen bonds with the phosphate group in Shc-like domains. D) Non-phosphorylated Tyr is recognized by hydrogen-bonding to one Gly and van der Waals interactions with a His in Dab-1 protein, but this is not a general recognition mode for Dab-like domains. Instead, these domains mainly interact with their targets through hydrogen-bonding and hydrophobic contacts to amino acids on their β5 strand. (Image A was adapted from reference [52]. Images B, C and D were reproduced from reference [51].)

1.2.2.3. Conserved domain 2 (C2)

Originally recognized as calcium-dependent phospholipid binders involved in signal transduction and vesicular trafficking, C2 domains have shown their ability to bind pTyr-containing protein motifs as well [54, 55]. These domains consist of approximately 130 amino acids structured into an eight-stranded β-sandwich. The crystal structures of two Protein kinases C (PKCδ and PKCθ) C2 domains in complex with an optimal sequence selected from peptide library screens were solved, revealing the key elements for pTyr recognition (Figure 1.11A) [54, 55]. C2 domains selectively bind to the (Tyr/Phe)-(Ser/Ala)-(Val/Ile)-pTyr-(Gln/Arg)-X-(Tyr/Phe)-X motif, where X is any amino acid. The peptide adopts an elongated conformation across one end of one β-sheet towards the other β-sheet

[54]. pTyr lies in a basic pocket and coordinates via hydrogen-bonding to the side chain and main chain NH group of one His residue, ring-stacking interactions between the phenyl group of pTyr and the same His residue, and electrostatic interactions with the side chains of one Arg and one Lys (Figure 1.11B). Other highly conserved residues further contribute for peptide recognition and binding affinity, through the establishment of side and main chain direct or water-mediated contacts [54, 55].

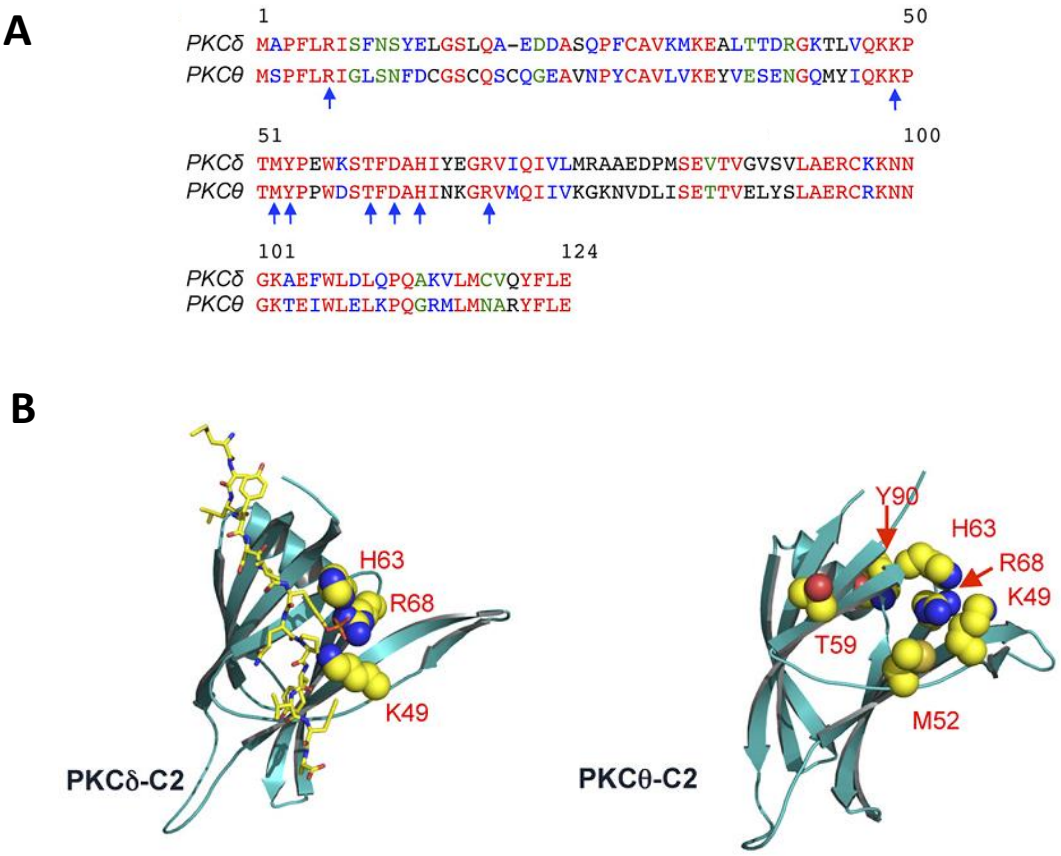


Figure 1.11 – C2 domains as pTyr-binding modules. A) Sequence homology between PKCδ and PKCθ C2 domains. Identical, highly similar, similar and non-conserved residues are coloured red, blue, green and black, respectively. Blue arrows indicate residues that are fundamental for peptide recognition. B) C2 domains of PKCδ and PKCθ form a β-sandwich composed of two four-stranded β-sheets (in light blue). The peptide (in stick representation) binds transversally to these β-sheets. The side chains of the amino acids involved in pTyr recognition are in space-filling representation and marked. (Reproduced from reference [55].)

1.3. Phosphoprotein and phosphopeptide enrichment

Phosphoproteomics is the systematic identification and characterization of phosphoproteins, usually by combining mass spectrometry (MS) strategies with phospho-specific enrichment methods. MS is the workhorse for the detection of phosphosites within a single protein or a complex protein mixture. Unfortunately, because of a low phosphorylation stoichiometry, a diversity of

phosphorylated protein forms and the transient nature of phosphorylation states, the amount of phosphorylated species in biological samples and protein digests is low. Thus, the enrichment of phosphoproteins and phosphopeptides *prior* to MS analysis is a mandatory step [2, 56]. This chapter highlights current and emerging platforms for phosphopeptide enrichment *prior* to MS analysis (Figure 1.12), because several recent reviews have successfully focused on MS analysis alone [57, 58]. For simplicity, the enrichment processes are described for phosphopeptides, as most phosphoprotein containing samples are enzymatically digested before further MS analysis.

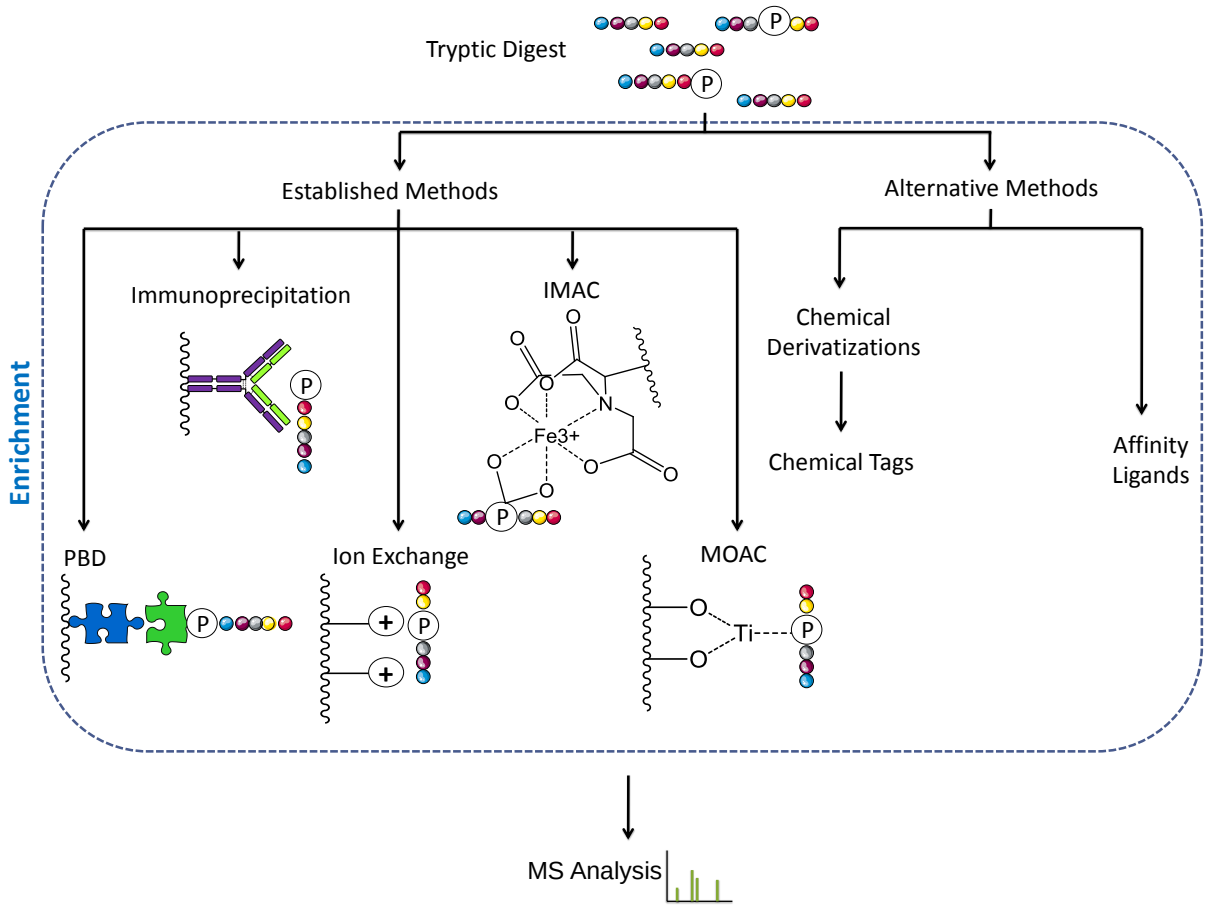


Figure 1.12 – Platforms for phosphopeptide enrichment. Established methods include immunoprecipitation, ion exchange chromatography (either Strong Cation Exchange (SCX) or Strong Anion Exchange (SAX)), Immobilized Metal Affinity Chromatography (IMAC), Metal Oxide Affinity Chromatography (MOAC) and, more recently, Phosphoprotein-binding Domains (PBDs). Novel tag-receptor systems and phosphate-affinity ligands have been developed as alternative methodologies. These methods can be further combined with enrichment platforms, as magnetic adsorbents and/or functionalized MALDI plates, before MS analysis.

1.3.1. Established strategies for phosphopeptide capture

Established approaches for the enrichment of phosphopeptides from complex biological samples involve chromatographic steps based on affinity, metal ion and ion exchange interactions (Figure 1.12). The platforms described in this section are those whose applications have been extended beyond the proof of concept with simple protein digests, and in which large-scale proteomic efforts using tissue samples have been implemented.

Affinity purification requires the coupling of an affinity ligand to a beaded matrix. In immunoaffinity chromatography, antibodies are utilized as affinity reagents, and therefore, this approach is limited by the commercial availability of antibodies against phosphorylated Ser, Tyr and Thr residues. The low specificity of antiphosphoserine and antiphosphothreonine antibodies has limited their widespread application as opposed to antiphosphotyrosine antibodies. Antiphosphotyrosine antibodies are more specific, reliable, and far more frequently used, being able to detect multiple Tyr phosphorylations on a variety of protein substrates [56, 59].

PDBs have also been used as ligands in affinity chromatography to target phosphorylated species [60]. Currently, there are 17 commercially available kits, based on different naturally occurring SH2 domains, for the selective capture of phosphotyrosine (pTyr) (Table 1.1).

Biological affinity reagents are expensive and usually labile; therefore, other strategies for the selective recognition of phosphorylated species have been explored using more durable and low-cost purification matrices.

Immobilized metal affinity chromatography (IMAC) and metal oxide affinity chromatography (MOAC) are widely used for phosphopeptide enrichment. In IMAC, metal ions (Fe^{3+}, Al^{3+}, Ga^{3+}, Zr^{4+}) are chelated to a stationary phase, such as silica, agarose beads or monoliths, via nitriloacetic acid (NTA), imidodiacetic acid (IDA) or tris-(carboxymethyl)-ethylenediamine ligands [61].

In MOAC, metal oxides, commonly titanium dioxide - TiO_2, are the anchoring point to entrap the phosphopeptide through the formation of multi-dentate bonds. Metal oxides can also be regarded as Lewis acids, where an ion exchange mechanism may be involved in the phosphate-metal oxide interaction. Several factors can influence the recognition of the phosphorylated species, such as synthesis conditions and the nature of the metal oxides, as recently reviewed [62]. Both IMAC and MOAC bind singly and multi-phosphorylated peptides. However, IMAC columns present stronger avidity towards multi-phosphorylated peptides. The multi-phosphorylated peptides present high binding affinity to MOAC supports, which makes elution difficult, and hence mono-phosphorylated species are more easily identifiable with MOAC [2, 63]. Additionally, a limitation in both IMAC and

MOAC, is the frequent nonspecific binding of peptides containing acidic amino acids (e.g. Glu and Asp) [2, 56]. Very recently a new technique, coined SIMAC (sequential elution from IMAC), has been reported as an improved version and combination of IMAC and TiO_2 [2].

Ion exchange chromatography represents another useful tool for the prefractionation of phosphopeptides *prior* to enrichment. In strong cation exchange (SCX) chromatography, acidic conditions promote the interaction between positively charged tryptic peptides and the negatively charged functional groups on the SCX particles. In appropriate acidic conditions, carboxyl groups will be protonated, and phosphopeptides will retain negative charges on phosphate groups. Phosphopeptides will be less positively charged, and therefore elute in earlier fractions than non-phosphorylated species [64]. Although the amount of contaminants co-adsorbing to the ion-exchange resin is still considerable, this type of adsorbent is usually combined with other enrichment methods, such as IMAC or MOAC [2, 56].

In anion exchange chromatography, phosphorylated species are typically more retained than their non-phosphorylated counterparts. However, acidic peptides bind strongly to the resin and the subsequent recovery of phosphopeptides is still a difficult task. Recently, a strong-anion exchange (SAX) capillary monolith was coupled with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) using buffers with pH values compatible with protein digestion and preparation of the MALDI matrix solution; phosphopeptides were detected at concentrations down to the attomole level [65].

Table 1.1 – Commercially available phosphoprotein and phosphopeptide enrichment kits.

Company	Kits
Calbiochem (http://www.merckmillipore.com/)	ProteoExtract® Phosphopeptide Enrichment TiO_2 Kit ProteoExtract® Phosphopeptide Capture Kit
Clontech (http://www.clontech.com/)	BD™ Phosphoprotein Enrichment Kit TALON® PMAC Magnetic Phospho Enrichment Kit
EMD4Biosciences (http://www.emdchemicals.com/)	Calbiochem® Phosphoprotein Enrichment Kit ProteoExtract® Phosphopeptide Enrichment TiO2 Kit
GE Healthcare (http://www.gelifesciences.com/)	TiO_2 Mag Sepharose™
GL Sciences (http://www.glsciences.com/)	Titansphere Phos-TiO Kit
Glygen (http://www.glysci.com/)	Phosphopeptide Enrichment NuTip Phosphopeptide Enrichment TopTip Phosphopeptide Enrichment Lab-in-a-Plate Phosphopeptide Enrichment Lab-in-a-Plate Flow-Thru Plate
Invitrogen (http://www.invitrogen.com/)	Captivate™ Microscale Phosphopeptide Isolation Kit Pro-Q® Diamond Phosphoprotein Enrichment Kit
Komabiotech (http://www.komabiotech.com/)	EzWay™ Phos-Pep Enrichment Kit EzWay™ Phos-Pro Enrichment Kit
Millipore (http://www.millipore.com/)	ZipTip with 0.6 µL metal chelate resin
Perkin Elmer (www.perkinelmer.com)	Phos-trap™ Enrichment Kit
Phos-Tag™ (http://www.phos-tag.com)	Phos-tag™ Agarose
Qiagen (http://www.qiagen.com/)	Phosphoprotein Purification Cartridge Phosphoprotein Purification Kit IMAC Mass Spec Focus Chips
Sigma-Aldrich (http://www.sigmaaldrich.com/)	Phos-Select™ Gallium Silica Spin Column Kit Phos-Select™ High Capacity Iron Coated Plate Phos-Select™ Iron Affinity Gel Supel-Tips Zr Pipette Tips Supel-Tips Ti Pipette Tips
Thermo Scientific (http://www.piercenet.com/)	Phosphoprotein Enrichment Kit SH2 Domain Phosphotyrosine Capture Kits (17 kits) Ga-IDA Phosphopeptide Enrichment Kit Fe-NTA Phosphopeptide Enrichment Kit Magnetic Titanium Dioxide Phosphopeptide Enrichment Kit TiO_2 Phosphopeptide Enrichment and Clean-up Kit
ZirChrom (www.zirchrom.com)	Titanium Dioxide SPE Tips Zirconium Dioxide SPE Tips

1.3.2. Combining traditional strategies with novel solid supports

1.3.2.1. On-Plate enrichment with MALDI-TOF MS

The detection of phosphorylated peptide fragments by MS is hampered by their low abundance, low ionization efficiencies, and by the ion suppression effect of the non-phosphopeptides. When positive ion mode is used, the negative charges on phosphate groups decrease the net charge of the ionized peptide. Moreover, phosphopeptides are usually less hydrophobic, and present low isoelectric point (pI) values and low-gas phase basicities, which makes them difficult to analyze by electrospray ionization mass spectrometry (ESI MS) or MALDI-TOF MS [66]. Negative ion mode is commonly used to circumvent the problem, because phosphopeptides present increased signal intensities. However, the determination of phosphorylation sites is challenging because negative polarity MS/MS spectra are difficult to interpret. One alternative would be to perform the detection of phosphopeptides by MS using negative ion mode, and then switch the polarity to positive ion mode when determining the phosphorylation site by MS/MS. This process represents an unpopular option because it is time-consuming and leads to decreased scanning rates [56]. Reducing sample complexity by applying phospho-enrichment methods will reduce the ion suppression effect. Therefore, the use of simplified protocols for capture, concentration and analysis of protein digests in a single step is extremely valuable. A single step process increases signal-to-noise (S/N) ratios, minimizes sample loss, and improves the detection of target molecules. ESI mass spectrometers are popular in protein phosphorylation analysis, because they can be easily coupled to separation techniques such as the commonly used liquid chromatography (LC). Until recently, the utilization of MALDI-TOF MS required a preceding enrichment step, utilizing methods presented in the previous section. It is now possible to integrate the process of enrichment and analysis through modification of MALDI plates with a variety of materials, where IMAC and MOAC functionalities remain the preferred option. The process consists of adding the sample to a functionalized MALDI probe, rinsing the sample on the probe, and finally adding the matrix [67].

Enhanced MALDI signals can be obtained by using patterned polycationic gold-coated silicon, such as poly(acrylic acid) (PAA) derivatized with polyethylenimine or complexed with Fe^{3+} [68]. Coupling an NTA derivative to the PAA surface, before complexation with Fe^{3+}, significantly improves the selectivity of the probe. Moreover, detection can be further enhanced by using 2',4',6'-trihydroxyacetophenone (THAP) mixed with diammonium hydrogen citrate (DAHC) as the MALDI matrix [69]. THAP/DAHC can be used either in positive or negative ion mode analysis, and increases

the S/N ratio for phosphopeptides by an order of magnitude, compared to the more commonly employed α-cyano-4-hydroxycinnamic acid (α-CHCA) matrix. The use of THAP/DAHC presents three main advantages: (i) reduction of sodium contaminants in the sample solution, because di-sodium citrate precipitates are formed during the co-crystallization procedure; (ii) enhancement of phosphopeptide ionization during positive ion mode MS, as a result of the formation of complexes between ammonium cations and phosphate groups; and (iii) detection of both phosphorylated and non-phosphorylated peptides as protonated molecules, because of the loss of ammonia during MS analysis [70].

Another option to target phosphopeptides makes use of NTA functional groups immobilized on a gold surface via a 1,8-octanedithiol linkage and followed by loading Fe^{3+} or Ga^{3+} ions on the surface. The procedure briefly consists of the immersion of gold substrates first in an ethanolic solution of 1,8-octanedithiol, which forms a self-assembled monolayer (SAM) on gold, followed by a solution of N-[5-(3-maleimidopropylamido)-1-carboxypentyl]-iminodiacetic acid, and finally a $FeCl_3$ or $Ga(NO_3)_3$ solution. Weak acidic conditions (pH 4-5) are required for the optimal enrichment of phosphorylated species. Moreover, Ga^{3+}-NTA surfaces seem to present better results in terms of reproducibility than Fe^{3+}-NTA surfaces. Although these SAMs offer a good performance in the on-probe enrichment of phosphopeptides, further optimization is needed with crude samples [71].

An IDA derivative, IDA-1,2-epoxy-9-decene, immobilized on the surface of a porous silicon wafer through a photochemical reaction, and then grafted with Fe^{3+}, is able to trap mono- and tetra-phosphopeptides specifically from tryptic digests of β-casein. 2,5-Dihydroxybenzoic acid (2,5-DHB) containing 1% phosphoric acid (H_3PO_4) was selected as the MALDI matrix, because it performed better than α-CHCA, probably because 2,5-DHB is also a metal chelating agent [72, 73]. The addition of H_3PO_4 to the matrix is known to enhance phosphopeptide ion signals during MALDI-TOF MS analysis [73].

Universal metal-chelating plastic MALDI chips for signal enhancement of phosphopeptides have also been developed. The chips are fabricated by the copolymerization of acrylic acid N-hydroxysuccinimide ester and methyl methacrylate monomers in a 1:13.3 molar ratio under UV exposure in a sandwich mold (glass plate, spacer and silicon master). The surface is activated with N,N-bis(carboxymethyl)-L-lysine hydrate, before metal chelation. Chelated hard metal ions (e.g. Fe^{3+}, Ga^{3+}, Al^{3+}) enrich phosphopeptides at low pH values, with Ga^{3+} being the most selective. Ga^{3+}-MALDI chips have been evaluated in terms of signal intensities, and are fully comparable to the commercially available PHOS-Select Iron Affinity Gel from Sigma for phosphopeptide analysis (Table

1.1). Intermediate metal ions, such as Ni^{2+} and Cu^{2+}, have demonstrated their efficiency to bind whole phosphorylated proteins at higher pH values [72].

MALDI plates modified with high-capacity polymer brushes are a promising alternative to conventional methodologies. Gold substrates functionalized with poly(2-hydroxyethyl methacrylate) (PHEMA) brushes and derivatized with Fe^{3+}-NTA, present a phosphopeptide binding capacity of 0.6 $\mu g/cm^2$, and a detection limit for tryptic peptides derived from β-casein of < 15 fmol. Fe^{3+}-NTA-PHEMA-modified plates exhibited phosphopeptide recoveries of around 70%, similar to the commercially available TiO_2 NuTips, and even higher than $ZipTip_{MC}$ pipet tips, ZrO_2 NuTips, and Qiagen IMAC Mass Spec Focus Chips (Table 1.1) [74]. These PHEMA-NTA-Fe^{3+} brushes might be enclosed by a hydrophobic poly(dimethylsiloxane) layer, reducing the spot diameter around 8 times, and subsequently lowering the sample volume and the detection limit by fivefold [67].

The approaches described above are based on IMAC-functionalized MALDI plates, therefore presenting the drawbacks related to conventional IMAC methods, such as the nonspecific binding of acidic peptides. A more sensitive alternative uses a zirconium phosphonate-modified porous silicon surface (ZrP-pSi), enabling the detection of phosphopeptides from the tryptic digestion of 2 fmol β-casein. The low detection limit allows for the identification of phosphorylated species even when a mixture of β-casein and bovine serum albumin in a molar ratio of 1:100 is used, whereas IMAC-based approaches lose specificity when high concentrations of non-phosphopeptides are present. However, multi-phosphorylated peptides are preferentially enriched relative to singly phosphorylated peptides using this approach, most probably because of a stronger interaction with the ZrP-pSi surface [75].

1.3.2.2. Magnetic phospho-fishing

During the past decade, superparamagnetic particles have been the focus of intense research, especially in the fields of bioseparation and biomedicine. They are non-porous adsorbents, easily manipulated by an external magnetic field, and present a high surface/volume ratio, leading to high binding capacities and fast binding kinetics [76]. The scope of application of magnetic particles for biotechnological separations is wide, but it is only relatively recently that magnetic materials have begun to be used for enrichment and identification of phosphopeptides. A variety of methods, which merge the fundamentals of metal ion affinity chromatography with magnetic separation have appeared in the past decade, and magnetic beads for phosphoproteomics analysis are already commercially available (Table 1.1).

In view of the success of using metal oxide coated adsorbents for the selective enrichment of phosphorylated species, alumina-coated iron oxide magnetic particles (Fe_3O_4@Al_2O_3) are able to concentrate phosphopeptides selectively from tryptic digests, with a capacity of 60 μg of phosphopeptides/g of particles. Specifically, Fe_3O_4@Al_2O_3 is mixed with an acidified tryptic digest solution containing phosphorylated species. The acidic conditions prevent the binding of non-phosphopeptides containing acidic amino acid residues. After washing with a solution of 0.15% trifluoroacetic acid (TFA) in acetonitrile (ACN)/water, the peptide-loaded Fe_3O_4@Al_2O_3 particles are mixed with a MALDI matrix (2,5-DHB/ 1% H_3PO_4), and deposited on the MALDI plate using the dried droplet method. The entire process, including enrichment step and MALDI-TOF MS analysis, is extremely fast taking only ≈5 min [77]. These particles provide higher selectivity towards phosphopeptides than Fe^{3+}-immobilized magnetic silica microspheres, commercial Fe^{3+}–IMAC resins, and TiO_2 beads [78]. Moreover, the interference of the particles in the laser desorption/ionization process is negligible, in view of the small size of the particles, which do not hinder signal intensities, resolution or mass accuracy of the mass spectra [79].

Zirconia-coated magnetic nanoparticles (Fe_3O_4@ZrO_2) are also applied in integrated phosphopeptide capturing and digestion. The protein-Fe_3O_4@ZrO_2 conjugates are heated in a microwave oven in the presence of trypsin, and digestion time was reduced to approximately 1 min. As magnetic nanoparticles absorb microwave radiation, the heating rate of the digestion solution is accelerated, and proteins attached to Fe_3O_4@ZrO_2 denature, becoming more vulnerable to enzymatic digestion [80].

Recently, other metal oxides, such as titanium dioxide (TiO_2) [81], gallium trioxide (Ga_2O_3) [82], tin dioxide (SnO_2) [83], tantalum pentoxide (Ta_2O_5) [84], niobium pentoxide (Nb_2O_5) [85], and zinc oxide (ZnO) [86] have also been used as coating materials for magnetic particles. Coating Fe_3O_4 nanoparticles with two metal oxides, TiO_2 and ZrO_2, allowed the capture of a larger amount of phosphopeptides when compared to Fe_3O_4@ZrO_2 or Fe_3O_4@TiO_2 alone, and efficiently enriched both mono- and multi-phosphorylated peptides [87].

Table 1.2 presents the detection limits for phosphopeptides, mostly obtained from tryptic digests of caseins and enriched using magnetic nanoparticles with different coatings. Although the MS analysis referred to in Table 1.2 were all conducted in MALDI-TOF MS instruments in positive ion mode, several factors might have influenced the sensitivity of the analysis. These include the digest and enrichment conditions, the type of mass spectrometer, laser or matrix utilized, and the maintenance conditions of the instrument.

As an alternative to regular ZrO_2, a novel strategy using zirconium phosphonate (Zr^{4+}-PO_3^{2-}) as a coating shell has been proposed. The strong multi-coordination effect between Zr^{4+}-PO_3^{2-} and phosphoric species provides a more selective and sensitive support. The process can be easily combined with SCX separation, resulting in a high performance methodology to enrich phosphopeptides from complex real samples [88, 89].

Iron oxide magnetic particles have also been coated with silica and grafted with a silane coupling agent (3-glycidoxypropyltrimethoxysilane with iminodiacetic acid – GLYMO–IDA). Fe^{3+} or Ce^{4+} ions can be immobilized on this surface, resulting in high phosphate selectivity; Ce^{4+} is particularly selective [90, 91]. Using the same principle, magnetic particles are coated with NTA and Zr^{4+} and Ga^{3+} ions to enrich phosphopeptides from complex real samples. Fe_3O_4/NTA-Zr^{4+} offered a better trapping performance than Fe_3O_4/NTA-Ga^{3+}, which is probably related to the coordination number of both metals [92].

Mesoporous materials consist of ordered porous networks with controllable pore sizes, high surface areas, and large pore volumes. Titanium grafted magnetic mesoporous silica (Fe_3O_4@Ti-mSiO_2) has been successfully applied for phosphopeptide enrichment, by combining the affinity and size exclusion properties of the support [93].

Bacterial magnetic nanoparticles (BMP) present an interesting alternative to the regularly used synthetic magnetic particles, because they do not require any surface functionalization before the immobilization of the metal ion. BMPs of *Magnetospirillum* consist of Fe_3O_4 crystals coated with bacterial lipid membrane, which contains phospho- and sulfolipids for excellent metal binding affinity. Both Zr^{4+} and Fe^{3+} ions were immobilized on the surface of BMP of *Magnetospirillum*. Zr^{4+}-immobilized BMP favored the enrichment of multi-phosphopeptides, whereas Fe^{3+}-immobilized BMP preferred singly phosphorylated peptides. Fe^{3+}-immobilized BMP presents a lower detection limit (Table 1.2), probably because of the larger amount of Fe^{3+} attached to the BMP. Moreover, BMP themselves were able to enrich a specific phosphopeptide from an α-casein digest through weak interactions [94].

Rare earth metals present strong affinity towards phosphate moieties. γ-Fe_2O_3 magnetic nanoparticles coated with a shell of ammonium fluoride and lutetium fluoride (γ-Fe_2O_3@xNH4F.yLuF_3) have been reported to efficiently enrich samples of β-casein tryptic digest, nonfat milk tryptic digest and human serum in phosphorylated peptides [95].

Particles coated with an yttrium phosphate (YPO_4) hollow porous affinity shell (Fe_3O_4@hYPO_4) have also been successfully used in phosphopeptide enrichment [96]. Faster adsorption/ desorption dynamics and low non-specific binding can be obtained by capping Fe_3O_4 first with polyacrylate and then with a ultrathin YPO_4 shell (PA-Fe_3O_4@YPO_4) [97].

Table 1.2 – Magnetic particles with different coatings for the enrichment of phosphopeptides[2].

Core	Shell	Coupling Agent	Metal Ions	MALDI matrix	Tryptic digest	Sensitivity	Refs.
Fe_3O_4[a]	P(HEMA-GMA)[b]	IDA	Fe^{3+}/ Ga^{3+}	2,5-DHB + 1% H_3PO_4 in 0.1% TFA/ 30% ACN	-	-	[98]
Fe_3O_4	SiO_2	GLYMO–IDA[c]	Fe^{3+}	2,5-DHB + 1% H_3PO_4 in 10% MeOH	FLTEpYVATR[d]	5 fmol	[91]
					α-casein	20 fmol	
Fe_3O_4	SiO_2	GLYMO–IDA[c]	Ce^{4+}	2,5-DHB + 1% H_3PO_4 in 50% ACN	-	-	[90]
Fe_3O_4	SiO_2	NTA[e] derivative	Ga^{3+}/ Zr^{4+}	2,5-DHB + 1% H_3PO_4 in ACN/ H_2O (2/1)	Mixture of α- and β-caseins	50 fmol	[92]
Fe_3O_4	Al_2O_3	-	-	2,5-DHB + 1% H_3PO_4 in ACN/ H_2O (2/1)	α-casein	25 fmol	[77]
Fe_3O_4	TiO_2	-	-	2,5-DHB + 1% H_3PO_4	β-casein	50 fmol	[81]
Fe_3O_4@C[a,f]	SnO_2	-	-	2,5-DHB + 1% H_3PO_4 in 50% ACN	β-casein	80 fmol	[83]
Fe_3O_4@C[f]	Ga_2O_3	-	-	2,5-DHB + 1% H_3PO_4 in 50% ACN	β-casein	40 fmol	[82]
Fe_3O_4	ZrO_2	-	-	2,5-DHB + 0.5% H_3PO_4 in ACN/ H_2O (2/1)	β-casein	45 fmol	[80]
Fe_3O_4	SiO_2	PO_3^{2-}	Zr^{4+}	2,5-DHB + 1% H_3PO_4 in 70% ACN	β-casein	0.5 fmol	[89]

[2] [a] The analytes on MALDI target were the phosphopeptides coupled to particles, except for these where peptides were previously eluted from the magnetic beads; [b] P(HEMA-GMA): poly(2-hydroxyethyl methacrylate-co-glycidyl methacrylate); [c] GLYMO-IDA: 3-glycidoxypropyltrimethoxysilane (GLYMO) with iminodiacetic acid (IDA); [d] FLTEpYVATR: pure standard phosphopeptide; [e] NTA: Nitriloacetic acid; [f] Fe_3O_4@C: Carbon-coated Fe_3O_4; 2,5-DHB: 2,5-Dihydroxybenzoic acid; H_3PO_4: Phosphoric acid; ACN: Acetonitrile; MeOH: Methanol; TFA: Trifluoroacetic acid; CHCA: a-Cyano-4-hydroxycinnamic acid.

Table 1.2 (cont.) – Magnetic particles with different coatings for the enrichment of phosphopeptides [3].

Core	Shell	Coupling Agent	Metal Ions	MALDI matrix	Tryptic digest	Sensitivity	Refs.
Fe_3O_4	ZnO	-	-	2,5-DHB + CHCA + 0.5% H_3PO_4 in ACN/ H_2O (2/1)	β-casein	2.5 fmol	[86]
Fe_3O_4	C18[g]	-	-	2,5-DHB + 1% H_3PO_4 in 75% ACN	α-casein	10 fmol	[99]
					β-casein	1 fmol	
BMP[h]	-	-	Fe^{3+}	2,5-DHB + 1% H_3PO_4	β-casein	0.5 pmol	[94]
			Zr^{4+}			2 pmol	
γ-Fe_2O_3 [a]	xNH4F·yLuF3[i]	-	-	2,5-DHB + 1% H_3PO_4 in 50% ACN	β-casein	50 pmol	[95]
Fe_3O_4 [a]	Ti-$mSiO_2$[j]	-	-	2,5-DHB + 1% H_3PO_4 in 50% ACN	β-casein	10 fmol	[93]
Fe_3O_4[a]	YPO_4	-	-	2,5-DHB + 1% H_3PO_4 in 50% ACN	β-casein	10 fmol	[96]
PA-Fe_3O_4[a,k]	YPO_4	-	-	2,5-DHB + 1% H_3PO_4 in 50% ACN	β-casein	8 fmol	[97]

1.3.3. Proof-of-concept strategies for phosphopeptide enrichment

1.3.3.1. Chemical tags

The conversion of the phosphate moieties into different functional groups permits the introduction of tag-receptor systems that provide an efficient alternative for phosphopeptide enrichment (Figure 1.13). A variety of chemical derivatization procedures are available for tagging phosphate groups in proteins and peptides. These include, for example, the combination of β-elimination, Michael addition and proteolysis [100].

Based on the concept of isotope-coded affinity tags [101], a phosphoprotein isotope-coded affinity tag (PhIAT) has been introduced. Phosphate moieties are chemically replaced by isotope and biotin-labeled affinity tags to enrich and measure quantitatively O-phosphorylation states in proteins. The

[3] [a] The analytes on MALDI target were the phosphopeptides coupled to particles, except for these where peptides were previously eluted from the magnetic beads; [g] C18: octadecyltrimethoxysilane; [h] BMP: Bacterial Magnetic Particles; [i] $xNH_4F.yLuF_3$: ammonium fluoride and lutetium fluoride; [j] Ti-$mSiO_2$: titanium grafted mesoporous silica; [k] PA-Fe_3O_4: polyacrylate capped Fe_3O_4; 2,5-DHB: 2,5-Dihydroxybenzoic acid; H_3PO_4: Phosphoric acid; ACN: Acetonitrile.

process consists of the hydroxide ion-mediated β-elimination of the O-phosphate moiety and Michael addition of either the light or heavy isotopic version of 1,2-ethanedithiol (EDT). The isotope-labeled peptide is then biotinylated using (+)-biotinyl-iodoacetamidyl-3,6-dioxaoctanediamine, and isolated from non-phosphorylated species using avidin chromatography. This method not only reveals the phosphorylation sites within a protein, but also selectively enriches the phosphopeptide content from complex protein samples. However, there are some disadvantages: (i) simultaneous identification of hydrophobic peptides; (ii) nonspecific binding of samples containing endogenous biotin and biotin-binding proteins; (iii) formation of diastereomers (i.e., epimers) upon Michael addition of EDT; and (iv) the need to use harsh denaturating conditions during elution [102].

A simplified, more sensitive and efficient approach based on site-specific labeling and isolation of cysteinyl- peptides has been proposed [103]. Here, the biotin affinity tag of PhIAT is replaced by a phosphoprotein isotope-coded solid-phase tag (PhIST). The tag is comprised of a photosensitive linker covalently bound to aminopropyl glass beads, a leucine isotope-coded linker containing six ^{12}C and one ^{14}N (light) or six ^{13}C and one ^{15}N (heavy), and a thiolate-reactive group. The captured peptides are released easily by UV photocleavage of the photosensitive linker and analyzed by capillary liquid chromatography tandem mass spectrometry (LC-MS/MS) [104].

Chemical labeling of phosphorylation sites can also be achieved through an arginine-mimic chemical tag, guanidinoethanethiol (GET). pSer residues are converted into guanidinoethylcysteine (Gec) by β-elimination/Michael addition in the presence of a barium catalyst, which accelerates the efficiency of the reaction. Proteolytic digestion occurs at Gec because it is recognized as a trypsin cleavage site. The method is applied to gel electrophoresis and also for on-bead chemical labeling coupled to TiO_2 enrichment. This method has been shown to be highly reproducible and sensitive, because basic guanidine moiety increases the peak intensities for the GET-labeled tryptic peptides by MS [105, 106].

Alternatively, a fluorescent affinity tag (FAT) consisting of rhodamine conjugated to a cysteamine moiety, selectively modifies phosphorylated Ser and Thr residues in a one-pot reaction. Subsequent use of commercially available anti-rhodamine affinity columns allows for the successful enrichment of FAT-labeled proteins and peptides [107].

These chemical labeling methods rely on a β-elimination reaction and, therefore, present drawbacks associated with non-specific labeling of cysteines and O-glycosylated residues. However, blocking the sulfhydryl group of cysteines by alkylation or oxidation, and performing enzymatic deglycosylation can reduce non-specific labeling. In addition, these methods are not applicable to pTyr residues, because they cannot undergo β-elimination [100].

Phosphoramidate chemistry can be used in the derivatization of phosphopeptides with sulfhydryl groups that are subsequently linked to iodoacetyl groups immobilized on a solid support. However, this approach involves six chemical reactions, leading to considerable sample loss mainly during the O-methylesterification steps. Despite a low final yield of ≈20%, this method is highly selective, resulting in contaminant-free phosphopeptides, and is equally applicable in the identification of pSer, pThr and pTyr moieties [108].

Phosphorylated residues can also be converted to their corresponding methyl esters and then subjected to a one-pot reaction in the presence of N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide (EDC), imidazole and a soluble dendrimer [109]. Although this method allows for high recoveries of phosphopeptides, the EDC-mediated catalysis reaction is extremely slow and requires blocking all carboxylate groups of the peptides, which is not very feasible or scalable.

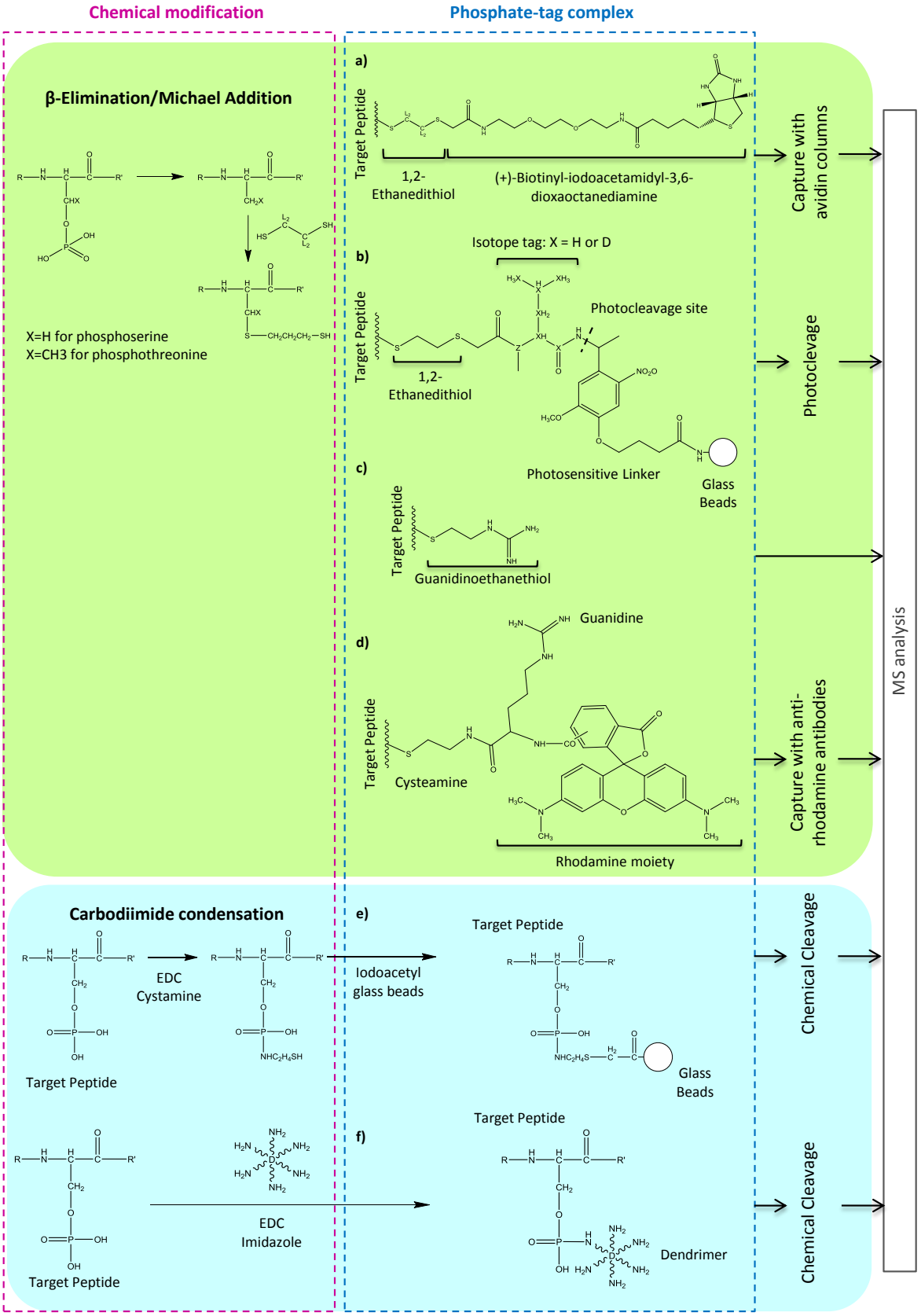


Figure 1.13 – Strategies for phosphopeptide enrichment and analysis using chemical tags. The most common chemical modifications rely on β-Elimination/Michael Addition and Carbodiimide Condensation.

Figure 1.13 (cont.) - (a) Phosphoprotein isotope-coded affinity tag (PhIAT): pSer and pThr containing proteins are isotopically labeled using light (L=H) or heavy (L=D) versions of 1,2-ethanedithiol (EDT) and biotinylated; (b) Phosphoprotein isotope-coded solid-phase tag (PhIST): the biotin moiety is replaced by aminopropyl beads covalently linked with a photosensitive linker, a stable isotope-coded leucine moiety, and a thiolate-reactive group; (c) Guanidinoethanethiol tag (GET): pSer containing proteins are derivatized through β-elimination and Michael addition of an arginine-mimic, guanidinoethanethiol. The guanidine moiety has excellent proton affinity, increasing the detection by MS; (d) Fluorescent affinity tag (FAT): pSer and pThr containing proteins undergo β-elimination/ Michael addition through cysteamine moiety of FAT, which also contains guanidine and rhodamine moieties, facilitating the enrichment via anti-rhodamine antibodies; Chemical tags based on (e) glass beads derivatized with iodoacetyl groups and (f) an aminated dendrimer, react by carbodiimide condensation with phosphorylated peptides. EDC: N-(3-dimethylaminopropyl)-N'-ethylcarbodiimide.

1.3.3.2. Phosphate-affinity ligands

Phosphate affinity tags rely on the chemical modification of phosphate groups with a molecule amenable to capture by affinity adsorption, whereas phosphate-affinity ligands specifically recognize phosphate moieties on peptides and proteins. The traditional affinity ligands for phosphopeptide enrichment referred to above (i.e. antibodies, SH2 domains and metal-chelating) are being replaced by alternative structures (Figure 1.14).

A new phosphate-binding molecule based on an alkoxide-bridged dinuclear metal complex – "Phos-tag" - has been reported and is commercially available (Table 1.1). This affinity ligand has passed the proof-of-concept stage and is one of the most promising for future implementation in large proteomic studies [110]. The complex was first reported as a dinuclear Zn(II) complex, 1,3-bis[bis(2-pyridinylmethyl)amino]-2-propanolato dizinc (II), but can also be complexed with manganese (Figure 1.14a). Phos-tag molecules have a vacancy on two metal ions, capturing phosphomonoester dianions bound to Ser, Thr, and Tyr residues [111-113]. This tag has been coupled to biotin [111], acrylamide [111] and agarose [114] and used in a range of phosphopeptide separation and identification procedures, such as affinity chromatography [114], electroblotting [115] and SDS-PAGE [116]. The phosphate-Phos-Tag complex, $ROPO_3^{2-}$ - $(Zn^{2+}$-Phos-Tag$)^{3+}$, has a characteristic mass shift of 581.1 Da and total charge that changes from -2 to +1, thereby facilitating the analysis of the phosphorylated compounds by MS [113]. As an alternative to the complex and laborious synthesis of Phos-Tag, dinuclear bispyridylmethylamine–tyrosine–acrylamide ligands have been prepared by a twofold Mannich reaction and used to enhance the separation of phosphoproteins in gel electrophoresis (Figure 1.14b) [117].

Other synthetic ligands include anthracene derivatives consisting of dinuclear Zinc(II)-dipicolylamine (Zn(II)-Dpa) receptors that exhibit high affinity to mono-phosphopeptides, and are based on the cooperative binding of both Zn(II)-Dpa sites (Figure 1.14c) [118]. These structures also show great

promise as identifiers of disease markers [119]. By coupling two Zn(II)-Dpa modules with 2,2'-bipyridine as a spacer in 4,4'-, 5,5'-, or 6,6'-substitution, high selectivity for multi-phosphopeptides has been achieved, as a result of multi-point interactions and intrapeptide crosslinking [120]. Rationally developed amphiphilic Zn^{2+}-cyclens have also proven to be efficient in the molecular recognition of phosphate moieties (Figure 1.14d) [121]. Enhanced phosphopeptide binding affinities have been observed under physiological conditions by coupling a cooperative second binding site that recognizes a carboxylate or imidazole amino acid side chain functionality [122, 123].

Molecularly imprinted polymers can also target phosphorylated species. A phosphate-selective molecularly imprinted polymer has been prepared using 1-allyl-2-urea as a functional monomer and a diphenylphosphate template. The imprinted polymer exhibited high binding ability and selectivity for phosphate in aqueous media, with a 70% phosphate recovery [124]. A phosphotyrosine-imprinted polymer receptor, using two urea-based monomers and an N,O-protected pTyr template, has been recently developed. These receptors have proved to be resistant, inexpensive, reusable, pTyr selective, and present binding constants for the amino acid template of around 10^7 M^{-1} in aqueous-rich solvents, comparable to the affinity constants of pTyr specific antibodies [125, 126].

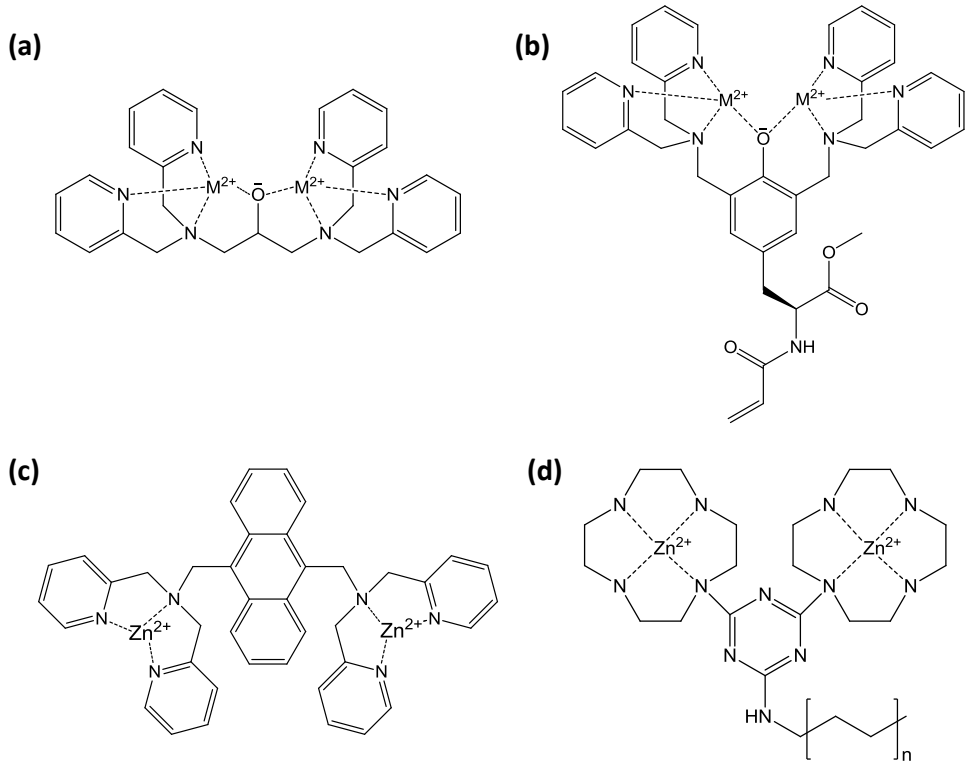


Figure 1.14 – Structure of phosphate-affinity ligands. (a) Phos-tag ligands consist of 1,3-bis[bis(2-pyridinylmethyl)amino]-2-propanolato dimetal(II) complex. (b) Dinuclear bispyridylmethylamine–tyrosine–acrylamide ligands. (c) Anthracene derivatives consisting of dinuclear zinc(II)-dipicolylamine receptors. (d) bis(Zn(II)-cyclen) ligands. M^{2+} is either Zn^{2+} or Mn^{2+}.

1.4. Perspectives

The success of phosphoproteomic studies results from a combination of efficient and selective phosphopeptide enrichment tools, as well as sensitive analytical techniques. While the analytical techniques are massively undertaken by a plethora of MS protocols, this chapter has focused on the current and future alternative methods for the enrichment of phosphorylated species. Immunoaffinity, metal ion and ion exchange interactions, coupled to chromatographic matrices, remain the obvious choice for large proteomic efforts. Nonetheless, there is an enormous interest in the search for more selective, low cost and integrated platforms for phosphopeptide enrichment [127]. In recent years, a substantial armory of enrichment methods for phosphoproteomic studies has been reported. Most described methods rely on the chelation between metal ions and the phosphate groups of proteins, or on the selective chemical modification of the phosphate moieties with functional groups amenable for detection and purification purposes. However, the high functional and structural diversity of phosphorylated proteins and peptides renders these techniques less specific. The identification of molecular recognition elements is fundamental for the design of context-independent, novel, affinity-molecules towards phosphorylated proteins and peptides [128]. Several international consortia are being established to identify systematically large collections of quality and validated affinity reagents against the human proteome [129, 130]. Antibodies and their engineered forms are the typical affinity reagents, but engineered proteins, nucleic acid aptamers and unnatural skeletons represent viable alternatives for this highly challenging biological target.

1.5. Bibliography

1. Carrascal, M., D. Ovelleiro, V. Casas, M. Gay, and J. Abian, *Phosphorylation analysis of primary human T lymphocytes using sequential IMAC and titanium oxide enrichment.* Journal of Proteome Research, 2008. **7**(12): p. 5167-5176.
2. Thingholm, T.E., O.N. Jensen, and M.R. Larsen, *Analytical strategies for phosphoproteomics.* Proteomics, 2009. **9**(6): p. 1451-1468.
3. Harsha, H. and A. Pandey, *Phosphoproteomics in cancer.* Molecular Oncology, 2010. **4**(6): p. 482-495.
4. Gloeckner, C.J., K. Boldt, F. von Zweydorf, S. Helm, L. Wiesent, H. Sarioglu, and M. Ueffing, *Phosphopeptide analysis reveals two discrete clusters of phosphorylation in the N-terminus and the Roc domain of the Parkinson-disease associated protein kinase LRRK2.* Journal of Proteome Research, 2010. **9**(4): p. 1738-1745.
5. Hanger, D.P., B.H. Anderton, and W. Noble, *Tau phosphorylation: the therapeutic challenge for neurodegenerative disease.* Trends in Molecular Medicine, 2009. **15**(3): p. 112-119.
6. Schwarz, E. and S. Bahn, *Biomarker discovery in psychiatric disorders.* Electrophoresis, 2008. **29**(13): p. 2884-2890.

7. Yu, L.R., H.J. Issaq, and T.D. Veenstra, *Phosphoproteomics for the discovery of kinases as cancer biomarkers and drug targets.* PROTEOMICS-Clinical Applications, 2007. **1**(9): p. 1042-1057.
8. Li, J., B.L. Williams, L.F. Haire, M. Goldberg, E. Wilker, D. Durocher, M.B. Yaffe, S.P. Jackson, and S.J. Smerdon, *Structural and Functional Versatility of the FHA Domain in DNA-Damage Signaling by the Tumor Suppressor Kinase Chk2.* Molecular Cell, 2002. **9**(5): p. 1045-1054.
9. Qin, B.Y., B.M. Chacko, S.S. Lam, M.P. de Caestecker, J.J. Correia, and K. Lin, *Structural Basis of Smad1 Activation by Receptor Kinase Phosphorylation.* Molecular Cell, 2001. **8**(6): p. 1303-1312.
10. Obsil, T. and V. Obsilova, *Structural basis of 14-3-3 protein functions.* Seminars in Cell & Developmental Biology, 2011. **22**: p. 663-672.
11. Smerdon, S.J. and M.B. Yaffe (2010) *Chapter 72 - Recognition of Phospho-Serine/Threonine Phosphorylated Proteins by Phospho-Serine/Threonine-Binding Domains*, In R.A. Bradshaw and E.A. Dennis (Ed.), Handbook of Cell Signaling (2nd Edition), Academic Press, p. 539-550.
12. Bridges, D. and G.B. Moorhead, *14-3-3 proteins: a number of functions for a numbered protein.* Science Signaling, 2005. **2005**(296): p. re10.
13. Yaffe, M.B. and S.J. Smerdon, *PhosphoSerine/threonine binding domains: you can't pSERious?* Structure, 2001. **9**(3): p. R33-R38.
14. Liu, D., J. Bienkowska, C. Petosa, R.J. Collier, H. Fu, and R. Liddington, *Crystal structure of the zeta isoform of the 14-3-3 protein.* Nature, 1995. **376**(6536): p. 191-194.
15. Sluchanko, N.N. and N.B. Gusev, *Oligomeric structure of 14-3-3 protein: What do we know about monomers?* FEBS letters, 2012. **586**: p. 4249–4256.
16. Muslin, A.J., J.W. Tanner, P.M. Allen, and A.S. Shaw, *Interaction of 14-3-3 with Signaling Proteins Is Mediated by the Recognition of Phosphoserine.* Cell, 1996. **84**(6): p. 889-897.
17. Watts, F.Z. and N.C. Brissett, *Linking up and interacting with BRCT domains.* DNA Repair, 2010. **9**(2): p. 103-108.
18. Mohammad, D.H. and M.B. Yaffe, *14-3-3 proteins, FHA domains and BRCT domains in the DNA damage response.* DNA Repair, 2009. **8**(9): p. 1009-1017.
19. Williams, R.S., M.S. Lee, D.D. Hau, and J.N.M. Glover, *Structural basis of phosphopeptide recognition by the BRCT domain of BRCA1.* Nature Structural & Molecular Biology, 2004. **11**(6): p. 519-525.
20. Glover, J.M., *Insights into the molecular basis of human hereditary breast cancer from studies of the BRCA1 BRCT domain.* Familial Cancer, 2006. **5**(1): p. 89-93.
21. Manke, I.A., D.M. Lowery, A. Nguyen, and M.B. Yaffe, *BRCT repeats as phosphopeptide-binding modules involved in protein targeting.* Science, 2003. **302**(5645): p. 636-639.
22. Stucki, M., J.A. Clapperton, D. Mohammad, M.B. Yaffe, S.J. Smerdon, and S.P. Jackson, *MDC1 Directly Binds Phosphorylated Histone H2AX to Regulate Cellular Responses to DNA Double-Strand Breaks.* Cell, 2005. **123**(7): p. 1213-1226.
23. Campbell, S.J., R.A. Edwards, and J.N.M. Glover, *Comparison of the Structures and Peptide Binding Specificities of the BRCT Domains of MDC1 and BRCA1.* Structure, 2010. **18**(2): p. 167-176.
24. Reinhardt, H.C. and M.B. Yaffe, *Phospho-Ser/Thr-binding domains: navigating the cell cycle and DNA damage response.* Nature Reviews Molecular Cell Biology, 2013. **14**(9): p. 563-580.
25. Liang, X. and S.R. Van Doren, *Mechanistic Insights into Phosphoprotein-Binding FHA Domains.* Accounts of Chemical Research, 2008. **41**(8): p. 991-999.
26. Bell, D.W., J.M. Varley, T.E. Szydlo, D.H. Kang, D.C. Wahrer, K.E. Shannon, M. Lubratovich, S.J. Verselis, K.J. Isselbacher, J.F. Fraumeni, J.M. Birch, F.P. Li, J.E. Garber, and D.A. Haber, *Heterozygous germ line hCHK2 mutations in Li-Fraumeni syndrome.* Science, 1999. **286**: p. 2528–2531.

27. Durocher, D., I.A. Taylor, D. Sarbassova, L.F. Haire, S.L. Westcott, S.P. Jackson, S.J. Smerdon, and M.B. Yaffe, *The Molecular Basis of FHA Domain:Phosphopeptide Binding Specificity and Implications for Phospho-Dependent Signaling Mechanisms.* Molecular Cell, 2000. **6**(5): p. 1169-1182.
28. Chong, P.A., B. Ozdamar, J.L. Wrana, and J.D. Forman-Kay, *Disorder in a Target for the Smad2 Mad Homology 2 Domain and Its Implications for Binding and Specificity.* Journal of Biological Chemistry, 2004. **279**(39): p. 40707-40714.
29. Kamato, D., M.L. Burch, T.J. Piva, H.B. Rezaei, M.A. Rostam, S. Xu, W. Zheng, P.J. Little, and N. Osman, *Transforming growth factor-β signalling: Role and consequences of Smad linker region phosphorylation.* Cellular Signalling, 2013. **25**(10): p. 2017-2024.
30. Rezaei, H.B., D. Kamato, G. Ansari, N. Osman, and P.J. Little, *Cell biology of Smad2/3 linker region phosphorylation in vascular smooth muscle.* Clinical and Experimental Pharmacology and Physiology, 2012. **39**(8): p. 661-667.
31. Wu, J.-W., M. Hu, J. Chai, J. Seoane, M. Huse, C. Li, D.J. Rigotti, S. Kyin, T.W. Muir, R. Fairman, J. Massagué, and Y. Shi, *Crystal Structure of a Phosphorylated Smad2: Recognition of Phosphoserine by the MH2 Domain and Insights on Smad Function in TGF-β Signaling.* Molecular Cell, 2001. **8**(6): p. 1277-1289.
32. Chacko, B.M., B.Y. Qin, A. Tiwari, G. Shi, S. Lam, L.J. Hayward, M. de Caestecker, and K. Lin, *Structural Basis of Heteromeric Smad Protein Assembly in TGF-β Signaling.* Molecular Cell, 2004. **15**(5): p. 813-823.
33. Elia, A.E.H., P. Rellos, L.F. Haire, J.W. Chao, F.J. Ivins, K. Hoepker, D. Mohammad, L.C. Cantley, S.J. Smerdon, and M.B. Yaffe, *The Molecular Basis for Phosphodependent Substrate Targeting and Regulation of Plks by the Polo-Box Domain.* Cell, 2003. **115**(1): p. 83-95.
34. Cheng, K.-Y., E.D. Lowe, J. Sinclair, E.A. Nigg, and L.N. Johnson, *The crystal structure of the human polo-like kinase-1 polo box domain and its phospho-peptide complex.* EMBO Journal, 2003. **22**(21): p. 5757-5768.
35. de Cárcer, G., G. Manning, and M. Malumbres, *From Plk1 to Plk5: Functional evolution of polo-like kinases.* Cell Cycle, 2011. **10**(14): p. 2255-2262.
36. de Cárcer, G., B. Escobar, A.M. Higuero, L. García, A. Ansón, G. Pérez, M. Mollejo, G. Manning, B. Meléndez, and J. Abad-Rodríguez, *Plk5, a polo box domain-only protein with specific roles in neuron differentiation and glioblastoma suppression.* Molecular and Cellular Biology, 2011. **31**(6): p. 1225-1239.
37. Li, D. and R. Roberts, *Human Genome and Diseases: WD-repeat proteins: structure characteristics, biological function, and their involvement in human diseases.* Cellular and Molecular Life Sciences CMLS, 2001. **58**(14): p. 2085-2097.
38. Smith, T.F., C. Gaitatzes, K. Saxena, and E.J. Neer, *The WD repeat: a common architecture for diverse functions.* Trends in Biochemical Sciences, 1999. **24**(5): p. 181-185.
39. Orlicky, S., X. Tang, A. Willems, M. Tyers, and F. Sicheri, *Structural Basis for Phosphodependent Substrate Selection and Orientation by the SCFCdc4 Ubiquitin Ligase.* Cell, 2003. **112**(2): p. 243-256.
40. Wu, G., G. Xu, B.A. Schulman, P.D. Jeffrey, J.W. Harper, and N.P. Pavletich, *Structure of a β-TrCP1-Skp1-β-catenin complex: destruction motif binding and lysine specificity of the SCFβ-TrCP1 ubiquitin ligase.* Molecular Cell, 2003. **11**(6): p. 1445-1456.
41. Evrard-Todeschi, N., J. Gharbi-Benarous, G. Bertho, G. Coadou, S. Megy, R. Benarous, and J.-P. Girault, *NMR studies for identifying phosphopeptide ligands of the HIV-1 protein Vpu binding to the F-box protein β-TrCP.* Peptides, 2006. **27**(1): p. 194-210.
42. Kato, Y., K. Nagata, M. Takahashi, L. Lian, J.J. Herrero, M. Sudol, and M. Tanokura, *Common Mechanism of Ligand Recognition by Group II/III WW Domains REDEFINING THEIR*

FUNCTIONAL CLASSIFICATION. Journal of Biological Chemistry, 2004. **279**(30): p. 31833-31841.

43. Ingham, R.J., K. Colwill, C. Howard, S. Dettwiler, C.S. Lim, J. Yu, K. Hersi, J. Raaijmakers, G. Gish, and G. Mbamalu, *WW domains provide a platform for the assembly of multiprotein networks.* Molecular and Cellular Biology, 2005. **25**(16): p. 7092-7106.
44. Verdecia, M.A., M.E. Bowman, K.P. Lu, T. Hunter, and J.P. Noel, *Structural basis for phosphoserine-proline recognition by group IV WW domains.* Nature Structural & Molecular Biology, 2000. **7**(8): p. 639-643.
45. Yaffe, M.B., *Phosphotyrosine-binding domains in signal transduction.* Nature Reviews Molecular Cell Biology, 2002. **3**(3): p. 177-186.
46. Pawson, T., G.D. Gish, and P. Nash, *SH2 domains, interaction modules and cellular wiring.* Trends in Cell Biology, 2001. **11**(12): p. 504-511.
47. Pawson, T. and P. Nash, *Protein–protein interactions define specificity in signal transduction.* Genes & Development, 2000. **14**(9): p. 1027-1047.
48. Songyang, Z., S. Shoelson, J. McGlade, P. Olivier, T. Pawson, X. Bustelo, M. Barbacid, H. Sabe, H. Hanafusa, and T. Yi, *Specific motifs recognized by the SH2 domains of Csk, 3BP2, fps/fes, GRB-2, HCP, SHC, Syk, and Vav.* Molecular and Cellular Biology, 1994. **14**(4): p. 2777-2785.
49. Zhou, S., S.E. Shoelson, M. Chaudhuri, G. Gish, T. Pawson, W.G. Haser, F. King, T. Roberts, S. Ratnofsky, and R.J. Lechleider, *SH2 domains recognize specific phosphopeptide sequences.* Cell, 1993. **72**(5): p. 767-778.
50. Waksman, G., S.E. Shoelson, N. Pant, D. Cowburn, and J. Kuriyan, *Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: crystal structures of the complexed and peptide-free forms.* Cell, 1993. **72**(5): p. 779-790.
51. Uhlik, M.T., B. Temple, S. Bencharit, A.J. Kimple, D.P. Siderovski, and G.L. Johnson, *Structural and Evolutionary Division of Phosphotyrosine Binding (PTB) Domains.* Journal of Molecular Biology, 2005. **345**(1): p. 1-20.
52. Yan, K.S., M. Kuti, and M.-M. Zhou, *PTB or not PTB – that is the question.* FEBS Letters, 2002. **513**(1): p. 67-70.
53. Chen, L., C. Liu, F.C.F. Ko, N. Xu, I.O.-l. Ng, J.W.P. Yam, and G. Zhu, *Solution Structure of the Phosphotyrosine Binding (PTB) Domain of Human Tensin2 Protein in Complex with Deleted in Liver Cancer 1 (DLC1) Peptide Reveals a Novel Peptide Binding Mode.* Journal of Biological Chemistry, 2012. **287**(31): p. 26104-26114.
54. Benes, C.H., N. Wu, A.E.H. Elia, T. Dharia, L.C. Cantley, and S.P. Soltoff, *The C2 Domain of PKCδ Is a Phosphotyrosine Binding Domain.* Cell, 2005. **121**(2): p. 271-280.
55. Stahelin, R.V., K.-F. Kong, S. Raha, W. Tian, H.R. Melowic, K.E. Ward, D. Murray, A. Altman, and W. Cho, *Protein Kinase Cϑ C2 Domain Is a Phosphotyrosine Binding Module That Plays a Key Role in Its Activation.* Journal of Biological Chemistry, 2012. **287**(36): p. 30518-30528.
56. Paradela, A. and J.P. Albar, *Advances in the analysis of protein phosphorylation.* Journal of Proteome Research, 2008. **7**(5): p. 1809-1818.
57. Boersema, P.J., S. Mohammed, and A.J. Heck, *Phosphopeptide fragmentation and analysis by mass spectrometry.* Journal of Mass Spectrometry, 2009. **44**(6): p. 861-878.
58. Palumbo, A.M., S.A. Smith, C.L. Kalcic, M. Dantus, P.M. Stemmer, and G.E. Reid, *Tandem mass spectrometry strategies for phosphoproteome analysis.* Mass Spectrometry Reviews, 2011. **30**(4): p. 600-625.
59. Nita-Lazar, A., H. Saito-Benz, and F.M. White, *Quantitative phosphoproteomics by mass spectrometry: past, present, and future.* Proteomics, 2008. **8**(21): p. 4433-4443.
60. Machida, K., C.M. Thompson, K. Dierck, K. Jablonowski, S. Kärkkäinen, B. Liu, H. Zhang, P.D. Nash, D.K. Newman, and P. Nollau, *High-throughput phosphotyrosine profiling using SH2 domains.* Molecular Cell, 2007. **26**(6): p. 899-915.

61. Block, H., B. Maertens, A. Spriestersbach, N. Brinker, J. Kubicek, R. Fabis, J. Labahn, and F. Schäfer, *Immobilized-metal affinity chromatography (IMAC): a review.* Methods in Enzymology, 2009. **463**: p. 439-473.
62. Leitner, A., *Phosphopeptide enrichment using metal oxide affinity chromatography.* TrAC Trends in Analytical Chemistry, 2010. **29**(2): p. 177-185.
63. Ficarro, S.B., M.L. McCleland, P.T. Stukenberg, D.J. Burke, M.M. Ross, J. Shabanowitz, D.F. Hunt, and F.M. White, *Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae.* Nature Biotechnology, 2002. **20**(3): p. 301-305.
64. Villén, J. and S.P. Gygi, *The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry.* Nature Protocols, 2008. **3**(10): p. 1630-1638.
65. Dong, M., M. Wu, F. Wang, H. Qin, G. Han, J. Dong, R.a. Wu, M. Ye, Z. Liu, and H. Zou, *Coupling strong anion-exchange monolithic capillary with MALDI-TOF MS for sensitive detection of phosphopeptides in protein digest.* Analytical Chemistry, 2010. **82**(7): p. 2907-2915.
66. Xu, Y., L. Zhang, H. Lu, and P. Yang, *Mass spectrometry analysis of phosphopeptides after peptide carboxy group derivatization.* Analytical Chemistry, 2008. **80**(21): p. 8324-8328.
67. Wang, W.-H. and M.L. Bruening, *Phosphopeptide enrichment on functionalized polymer microspots for MALDI-MS analysis.* Analyst, 2009. **134**(3): p. 512-518.
68. Xu, Y., M.L. Bruening, and J.T. Watson, *Use of polymer-modified MALDI-MS probes to improve analyses of protein digests and DNA.* Analytical Chemistry, 2004. **76**(11): p. 3106-3111.
69. Dunn, J.D., J.T. Watson, and M.L. Bruening, *Detection of phosphopeptides using Fe (III)-nitrilotriacetate complexes immobilized on a MALDI plate.* Analytical Chemistry, 2006. **78**(5): p. 1574-1580.
70. Yang, X., H. Wu, T. Kobayashi, R.J. Solaro, and R.B. van Breemen, *Enhanced ionization of phosphorylated peptides during MALDI TOF mass spectrometry.* Analytical Chemistry, 2004. **76**(5): p. 1532-1536.
71. Shen, J., T. Ahmed, A. Vogt, J. Wang, J. Severin, R. Smith, S. Dorwin, R. Johnson, J. Harlan, and T. Holzman, *Preparation and characterization of nitrilotriacetic-acid-terminated self-assembled monolayers on gold surfaces for matrix-assisted laser desorption ionization-time of flight-mass spectrometry analysis of proteins and peptides.* Analytical Biochemistry, 2005. **345**(2): p. 258-269.
72. Ibáñez, A.J., A. Muck, and A. Svatoš, *Metal-chelating plastic MALDI (pMALDI) chips for the enhancement of phosphorylated-peptide/protein signals.* Journal of Proteome Research, 2007. **6**(9): p. 3842-3848.
73. Xu, S., H. Zhou, C. Pan, Y. Fu, Y. Zhang, X. Li, M. Ye, and H. Zou, *Iminodiacetic acid derivatized porous silicon as a matrix support for sample pretreatment and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry analysis.* Rapid Communications in Mass Spectrometry, 2006. **20**(11): p. 1769-1775.
74. Dunn, J.D., E.A. Igrisan, A.M. Palumbo, G.E. Reid, and M.L. Bruening, *Phosphopeptide enrichment using MALDI plates modified with high-capacity polymer brushes.* Analytical Chemistry, 2008. **80**(15): p. 5727-5735.
75. Zhou, H., S. Xu, M. Ye, S. Feng, C. Pan, X. Jiang, X. Li, G. Han, Y. Fu, and H. Zou, *Zirconium phosphonate-modified porous silicon for highly specific capture of phosphopeptides and MALDI-TOF MS analysis.* Journal of Proteome Research, 2006. **5**(9): p. 2431-2437.
76. Batalha, I.L., A. Hussain, and A. Roque, *Gum Arabic coated magnetic nanoparticles with affinity ligands specific for antibodies.* Journal of Molecular Recognition, 2010. **23**(5): p. 462-471.
77. Chen, C.-T., W.-Y. Chen, P.-J. Tsai, K.-Y. Chien, J.-S. Yu, and Y.-C. Chen, *Rapid Enrichment of Phosphopeptides and Phosphoproteins from Complex Samples Using Magnetic Particles*

Coated with Alumina as the Concentrating Probes for MALDI MS Analysis. Journal of Proteome Research, 2007. **6**(1): p. 316-325.

78. Li, Y., Y. Liu, J. Tang, H. Lin, N. Yao, X. Shen, C. Deng, P. Yang, and X. Zhang, *Fe3O4@ Al2O3 magnetic core-shell microspheres for rapid and highly specific capture of phosphopeptides with mass spectrometry analysis.* Journal of Chromatography A, 2007. **1172**(1): p. 57.
79. Chang, S.Y., N.-Y. Zheng, C.-S. Chen, C.-D. Chen, Y.-Y. Chen, and C. Wang, *Analysis of peptides and proteins affinity-bound to iron oxide nanoparticles by MALDI MS.* Journal of the American Society for Mass Spectrometry, 2007. **18**(5): p. 910-918.
80. Lo, C.-Y., W.-Y. Chen, C.-T. Chen, and Y.-C. Chen, *Rapid enrichment of phosphopeptides from tryptic digests of proteins using iron oxide nanocomposites of magnetic particles coated with zirconia as the concentrating probes.* Journal of Proteome Research, 2007. **6**(2): p. 887-893.
81. Chen, C.-T. and Y.-C. Chen, *Fe3O4/TiO2 core/shell nanoparticles as affinity probes for the analysis of phosphopeptides using TiO2 surface-assisted laser desorption/ionization mass spectrometry.* Analytical Chemistry, 2005. **77**(18): p. 5912-5919.
82. Li, Y., H. Lin, C. Deng, P. Yang, and X. Zhang, *Highly selective and rapid enrichment of phosphorylated peptides using gallium oxide-coated magnetic microspheres for MALDI-TOF-MS and nano-LC-ESI-MS/MS/MS analysis.* Proteomics, 2008. **8**(2): p. 238-249.
83. Qi, D., J. Lu, C. Deng, and X. Zhang, *Magnetically Responsive Fe3O4@ C@ SnO2 Core− Shell Microspheres: Synthesis, Characterization and Application in Phosphoproteomics.* The Journal of Physical Chemistry C, 2009. **113**(36): p. 15854-15861.
84. Lin, H.-Y., W.-Y. Chen, and Y.-C. Chen, *Iron oxide/tantalum oxide core–shell magnetic nanoparticle-based microwave-assisted extraction for phosphopeptide enrichment from complex samples for MALDI MS analysis.* Analytical and Bioanalytical Chemistry, 2009. **394**(8): p. 2129-2136.
85. Lin, H.-Y., W.-Y. Chen, and Y.-C. Chen, *Iron Oxide/Niobium Oxide CoreShell Magnetic Nanoparticle-Based Phosphopeptide Enrichment from Biological Samples for MALDI MS Analysis.* Journal of Biomedical Nanotechnology, 2009. **5**(2): p. 215-223.
86. Chen, W.-Y. and Y.-C. Chen, *Functional Fe3O4@ ZnO magnetic nanoparticle-assisted enrichment and enzymatic digestion of phosphoproteins from saliva.* Analytical and Bioanalytical Chemistry, 2010. **398**(5): p. 2049-2057.
87. Li, W., Q. Deng, G. Fang, Y. Chen, J. Zhan, and S. Wang, *Facile synthesis of Fe3O4@TiO2-ZrO2 and its application in phosphopeptide enrichment.* Journal of Materials Chemistry B, 2013. **1**(14): p. 1947-1961.
88. Wei, J., Y. Zhang, J. Wang, F. Tan, J. Liu, Y. Cai, and X. Qian, *Highly efficient enrichment of phosphopeptides by magnetic nanoparticles coated with zirconium phosphonate for phosphoproteome analysis.* Rapid Communications in Mass Spectrometry, 2008. **22**(7): p. 1069-1080.
89. Zhao, L., R.a. Wu, G. Han, H. Zhou, L. Ren, R. Tian, and H. Zou, *The highly selective capture of phosphopeptides by zirconium phosphonate-modified magnetic nanoparticles for phosphoproteome analysis.* Journal of the American Society for Mass Spectrometry, 2008. **19**(8): p. 1176-1186.
90. Li, Y., D. Qi, C. Deng, P. Yang, and X. Zhang, *Cerium ion-chelated magnetic silica microspheres for enrichment and direct determination of phosphopeptides by matrix-assisted laser desorption ionization mass spectrometry.* Journal of Proteome Research, 2008. **7**(4): p. 1767-1777.
91. Tan, F., Y. Zhang, W. Mi, J. Wang, J. Wei, Y. Cai, and X. Qian, *Enrichment of phosphopeptides by $Fe3^{+}$-immobilized magnetic nanoparticles for phosphoproteome analysis of the plasma membrane of mouse liver.* Journal of Proteome Research, 2008. **7**(3): p. 1078-1087.

92. Li, Y.-C., Y.-S. Lin, P.-J. Tsai, C.-T. Chen, W.-Y. Chen, and Y.-C. Chen, *Nitrilotriacetic acid-coated magnetic nanoparticles as affinity probes for enrichment of histidine-tagged proteins and phosphorylated peptides.* Analytical Chemistry, 2007. **79**(19): p. 7519-7525.
93. Li, X.-S., Y.-N. Pan, Y. Zhao, B.-F. Yuan, L. Guo, and Y.-Q. Feng, *Preparation of titanium-grafted magnetic mesoporous silica for the enrichment of endogenous serum phosphopeptides.* Journal of Chromatography A, 2013. **1315**: p. 61-69.
94. Huang, J., L. Guo, and L.-M. Zheng, *Rapid enrichment and determination of phosphopeptides using bacterial magnetic particles via both strong and weak interactions.* Analyst, 2010. **135**(3): p. 559-563.
95. Wang, Z.-G., G. Cheng, Y.-L. Liu, J.-L. Zhang, D.-H. Sun, and J.-Z. Ni, *Novel 3D flowerlike hierarchical γ-Fe2O3@xNH4FyLuF3 core-shell microspheres tailor-made by a phase transformation process for the capture of phosphopeptides.* Journal of Materials Chemistry B, 2013. **1**(37): p. 4845-4854.
96. Cheng, G., Y.-L. Liu, Z.-G. Wang, S.-M. Li, J.-L. Zhang, and J.-Z. Ni, *Yolk-shell magnetic microspheres with mesoporous yttrium phosphate shells for selective capture and identification of phosphopeptides.* Journal of Materials Chemistry B, 2013. **1**(30): p. 3661-3669.
97. Sun, Y. and H.-F. Wang, *Ultrathin-yttrium phosphate-shelled polyacrylate-ferriferrous oxide magnetic microspheres for rapid and selective enrichment of phosphopeptides.* Journal of Chromatography A, 2013. **1316**: p. 62-68.
98. Novotna, L., T. Emmerova, D. Horak, Z. Kucerova, and M. Ticha, *Iminodiacetic acid-modified magnetic poly (2-hydroxyethyl methacrylate)-based microspheres for phosphopeptide enrichment.* Journal of Chromatography A, 2010. **1217**(51): p. 8032-8040.
99. Hsiao, H.-H., H.-Y. Hsieh, C.-C. Chou, S.-Y. Lin, A.H.-J. Wang, and K.-H. Khoo, *Concerted experimental approach for sequential mapping of peptides and phosphopeptides using C18-functionalized magnetic nanoparticles.* Journal of Proteome Research, 2007. **6**(4): p. 1313-1324.
100. Knight, Z.A., B. Schilling, R.H. Row, D.M. Kenski, B.W. Gibson, and K.M. Shokat, *Phosphospecific proteolysis for mapping sites of protein phosphorylation.* Nature Biotechnology, 2003. **21**(9): p. 1047-1054.
101. Gygi, S.P., B. Rist, S.A. Gerber, F. Turecek, M.H. Gelb, and R. Aebersold, *Quantitative analysis of complex protein mixtures using isotope-coded affinity tags.* Nature Biotechnology, 1999. **17**(10): p. 994-999.
102. Goshe, M.B., T.D. Veenstra, E.A. Panisko, T.P. Conrads, N.H. Angell, and R.D. Smith, *Phosphoprotein isotope-coded affinity tags: application to the enrichment and identification of low-abundance phosphoproteins.* Analytical Chemistry, 2002. **74**(3): p. 607-616.
103. Zhou, H., J.A. Ranish, J.D. Watts, and R. Aebersold, *Quantitative proteome analysis by solid-phase isotope tagging and mass spectrometry.* Nature Biotechnology, 2002. **20**(5): p. 512-515.
104. Qian, W.-J., M.B. Goshe, D.G. Camp, L.-R. Yu, K. Tang, and R.D. Smith, *Phosphoprotein isotope-coded solid-phase tag approach for enrichment and quantitative analysis of phosphopeptides from complex mixtures.* Analytical Chemistry, 2003. **75**(20): p. 5441-5450.
105. Ahn, Y.H., E.S. Ji, K.H. Kwon, J.Y. Lee, K. Cho, J.Y. Kim, H.J. Kang, H.G. Kim, and J.S. Yoo, *Protein phosphorylation analysis by site-specific arginine-mimic labeling in gel electrophoresis and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry.* Analytical Biochemistry, 2007. **370**(1): p. 77-86.
106. Ahn, Y.H., E.S. Ji, J.Y. Lee, K. Cho, and J.S. Yoo, *Coupling of TiO2-mediated enrichment and on-bead guanidinoethanethiol labeling for effective phosphopeptide analysis by matrix-assisted*

laser desorption/ionization mass spectrometry. Rapid Communications in Mass Spectrometry, 2007. **21**(24): p. 3987-3994.

107. Stevens, S.M., A.Y. Chung, M.C. Chow, S.H. McClung, C.N. Strachan, A.C. Harmon, N.D. Denslow, and L. Prokai, *Enhancement of phosphoprotein analysis using a fluorescent affinity tag and mass spectrometry.* Rapid Communications in Mass Spectrometry, 2005. **19**(15): p. 2157-2162.
108. Zhou, H., J.D. Watts, and R. Aebersold, *A systematic approach to the analysis of protein phosphorylation.* Nature Biotechnology, 2001. **19**(4): p. 375-378.
109. Tao, W.A., B. Wollscheid, R. O'Brien, J.K. Eng, X.-j. Li, B. Bodenmiller, J.D. Watts, L. Hood, and R. Aebersold, *Quantitative phosphoproteome analysis using a dendrimer conjugation chemistry and tandem mass spectrometry.* Nature Methods, 2005. **2**(8): p. 591-598.
110. Yang, L., Z. Xue, Y. He, S. Sun, H. Chen, and L. Qi, *A Phos-tag-based approach reveals the extent of physiological endoplasmic reticulum stress.* PLoS One, 2010. **5**(7): p. e11621.
111. Kinoshita, E., E. Kinoshita-Kikuta, K. Takiyama, and T. Koike, *Phosphate-binding tag, a new tool to visualize phosphorylated proteins.* Molecular & Cellular Proteomics, 2006. **5**(4): p. 749-757.
112. Kinoshita, E., M. Takahashi, H. Takeda, M. Shiro, and T. Koike, *Recognition of phosphate monoester dianion by an alkoxide-bridged dinuclear zinc (II) complex.* Dalton Transactions, 2004(8): p. 1189-1193.
113. Takeda, H., A. Kawasaki, M. Takahashi, A. Yamada, and T. Koike, *Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of phosphorylated compounds using a novel phosphate capture molecule.* Rapid Communications in Mass Spectrometry, 2003. **17**(18): p. 2075-2081.
114. Kinoshita, E., A. Yamada, H. Takeda, E. Kinoshita-Kikuta, and T. Koike, *Novel immobilized zinc (II) affinity chromatography for phosphopeptides and phosphorylated proteins.* Journal of Separation Science, 2005. **28**(2): p. 155-162.
115. Kinoshita-Kikuta, E., E. Kinoshita, and T. Koike, *Phos-tag beads as an immunoblotting enhancer for selective detection of phosphoproteins in cell lysates.* Analytical Biochemistry, 2009. **389**(1): p. 83-85.
116. Kinoshita, E., E. Kinoshita-Kikuta, and T. Koike, *Separation and detection of large phosphoproteins using Phos-tag SDS-PAGE.* Nature Protocols, 2009. **4**(10): p. 1513-1521.
117. Dirscherl, G., M. Schwab, W. Seufert, and B. König, *Enhancing the separation of phosphorylated proteins in gel electrophoresis with dinuclear bispyridylmethylamine–tyrosine–acrylamide complexes.* Inorganica Chimica Acta, 2009. **362**(2): p. 537-542.
118. Ojida, A., Y. Mito-oka, M.-a. Inoue, and I. Hamachi, *First artificial receptors and chemosensors toward phosphorylated peptide in aqueous solution.* Journal of the American Chemical Society, 2002. **124**(22): p. 6256-6258.
119. Ojida, A., T. Sakamoto, M.-a. Inoue, S.-h. Fujishima, G. Lippens, and I. Hamachi, *Fluorescent BODIPY-based Zn (II) complex as a molecular probe for selective detection of neurofibrillary tangles in the brains of Alzheimer's disease patients.* Journal of the American Chemical Society, 2009. **131**(18): p. 6543-6548.
120. Ojida, A., M.-a. Inoue, Y. Mito-oka, and I. Hamachi, *Cross-linking strategy for molecular recognition and fluorescent sensing of a multi-phosphorylated peptide in aqueous solution.* Journal of the American Chemical Society, 2003. **125**(34): p. 10184-10185.
121. Turygin, D.S., M. Subat, O.A. Raitman, S.L. Selector, V.V. Arslanov, B. König, and M.A. Kalinina, *Two-dimensional arrays of amphiphilic Zn2+-cyclens for guided molecular recognition at interfaces.* Langmuir, 2007. **23**(5): p. 2517-2524.

122. Grauer, A., A. Riechers, S. Ritter, and B. König, *Synthetic receptors for the differentiation of phosphorylated peptides with nanomolar affinities.* Chemistry-A European Journal, 2008. **14**(29): p. 8922-8927.
123. Riechers, A., A. Grauer, S. Ritter, B. Sperl, T. Berg, and B. König, *Binding of phosphorylated peptides and inhibition of their interaction with disease-relevant human proteins by synthetic metal-chelate receptors.* Journal of Molecular Recognition, 2010. **23**(3): p. 329-334.
124. Kugimiya, A. and H. Takei, *Selectivity and recovery performance of phosphate-selective molecularly imprinted polymer.* Analytica Chimica Acta, 2008. **606**(2): p. 252-256.
125. Emgenbroich, M., C. Borrelli, S. Shinde, I. Lazraq, F. Vilela, A.J. Hall, J. Oxelbark, E. De Lorenzi, J. Courtois, and A. Simanova, *A Phosphotyrosine-imprinted Polymer Receptor for the Recognition of Tyrosine Phosphorylated Peptides.* Chemistry-A European Journal, 2008. **14**(31): p. 9516-9529.
126. Helling, S., S. Shinde, F. Brosseron, A. Schnabel, T. Müller, H.E. Meyer, K. Marcus, and B.r. Sellergren, *Ultratrace enrichment of tyrosine phosphorylated peptides on an imprinted polymer.* Analytical Chemistry, 2011. **83**(5): p. 1862-1865.
127. Björling, E., C. Lindskog, P. Oksvold, J. Linné, C. Kampf, S. Hober, M. Uhlén, and F. Pontén, *A web-based tool for in silico biomarker discovery based on tissue-specific protein profiles in normal and cancer tissues.* Molecular & Cellular Proteomics, 2008. **7**(5): p. 825-844.
128. Roque, A.C.A. and C.R. Lowe, *Lessons from nature: On the molecular recognition elements of the phosphoprotein binding-domains.* Biotechnology and Bioengineering, 2005. **91**(5): p. 546-555.
129. Stoevesandt, O. and M.J. Taussig, *Affinity reagent resources for human proteome detection: initiatives and perspectives.* Proteomics, 2007. **7**(16): p. 2738-2750.
130. Taussig, M.J., O. Stoevesandt, C.A. Borrebaeck, A.R. Bradbury, D. Cahill, C. Cambillau, A. De Daruvar, S. Dübel, J. Eichler, and R. Frank, *ProteomeBinders: planning a European resource of affinity reagents for analysis of the human proteome.* Nature Methods, 2007. **4**(1): p. 13-17.

Chapter 1 Platforms for enrichment of phosphorylated proteins and peptides in proteomics

CHAPTER 2

Combinatorial libraries of synthetic ligands for phosphorylated peptides

Summary

The structural study of phosphoprotein-binding domains (PBDs), and the identification of recognition elements, such as key amino acid residues and three-dimensional constraints, between these PBDs and their phosphorylated targets, inspires the development of tailor-made molecules able to selectively bind phosphorylated proteins and peptides. In this work, combinatorial libraries of solid-phase biomimetic ligands have been developed using three different molecular skeletons based on the triazine reaction, the Ugi reaction, and a tandem Petasis-Ugi reaction. The ligands were further screened against a small peptide comprising a consensus sequence (pSer-X-X-Phe, where pSer is phosphoserine and X is any amino acid) that specifically binds the Brca1 C-terminal (BRCT) domain. Lead ligands were selected based on their binding capacities and phosphopeptide enrichment efficiencies.

Chapter 2 Combinatorial libraries of synthetic ligands for phosphorylated peptides

2.1. Introduction

Affinity ligands are classified in two major classes: biospecific and pseudobiospecific. Biospecific ligands are the natural binding partners of a particular target. Pseudobiospecific ligands are chemically defined molecules, either biological (e.g. peptides, aptamers) or synthetic (e.g. hydrophobic, thiophilic, mixed-mode, meta-chelating ligands), which are able to specifically bind the same target. Biomimetic ligands are engineered molecules with improved features which belong to one of the subclasses of pseudobiospecific ligands (e.g. engineered protein domains, de novo designed ligands, etc) [1, 2]. This class of ligands has been developed to overcome drawbacks of their naturally-occurring templates, presenting high chemical resistance to cleaning-in-place and sterilization-in-place procedures, high scalability, high binding capacities and low production costs [1, 3]. They have been used for the purification of a myriad of target molecules, such as antibodies [4, 5], glycoenzymes [6], insulin precursors [7], and clotting factors [8], among others.

Ligand design has been led by four strategies: (i) use of natural binder as template; (ii) use of a molecule that is known to bind the target site as template; (iii) use the target molecule itself to study the interactions around the binding site; and (iv) examine and mimic the biological interactions between the natural ligand and the target [9]. The fourth strategy was the one followed in this work.

Besides the ligand *per se*, other considerations must be taken into account in order to obtain high affinity and target selectivity, such as the nature of the matrix and the chemistry used for ligand coupling. The matrix should establish as few non-specific interactions with the target molecule as possible, and care must be taken upon ligand immobilization in order to keep its functional moiety accessible for target binding. Moreover, the chemical bonds between the ligand and the matrix should remain stable during adsorption and elution steps [10].

The discovery and production of novel ligands for different targets is urged by massive breakthroughs in computational modelling, high-throughput screening methodologies and combinatorial synthesis approaches. Both biological and synthetic biomimetics have been successfully synthesized by combinatorial strategies, which include synthetic peptide and small chemical library synthesis, phage display for the discovery of biological ligands, among others. Different combinatorial methodologies and their respective advantages and drawbacks are beyond the scope and will not be extensively discussed here, but can be reviewed elsewhere [11-13].

The synthesis of large libraries of compounds directly attached to a solid support eliminates laborious intermediate purification steps, allowing the ease removal of excess reagents, by-products and solvents by a separation process, such as filtration [13]. Ligands synthesized in solution-phase, on the

other hand, are more easily characterized by mainstream techniques, such as nuclear magnetic resonance (NMR) and mass spectrometry (MS), but their post-synthesis matrix coupling remains a challenge.

One critical factor upon the design of these combinatorial libraries is to ensure molecular diversity. This can be achieved by introducing distinct appendages into a same molecular skeleton (appendage diversity); using chiral reagents (stereochemical diversity); and using different molecular skeletons (scaffold diversity) [14].

In the present work, three combinatorial libraries of fully synthetic ligands were built based on three chemical reactions: triazine, Ugi and a tandem Petasis-Ugi reaction.

The triazine reaction consists in a modified 'mix-and-split' procedure, where a cyanuric-chloride-modified matrix reacts by two subsequent nucleophilic substitutions (Figure 2.1) [15].

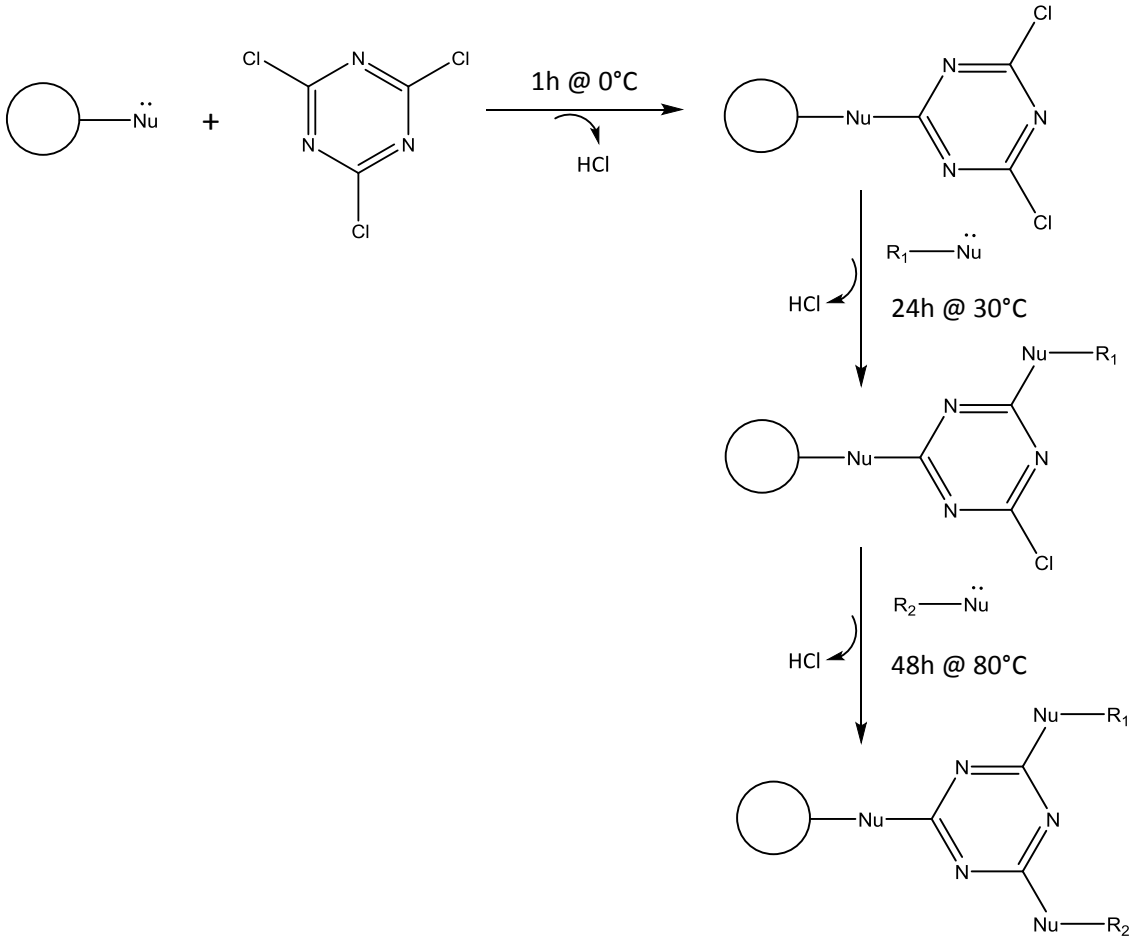


Figure 2.1 – Solid-phase synthesis of triazine-based scaffolds. A solid matrix functionalized with a nucleophile (denoted by Nü) readily reacts with cyanuric chloride at 0°C. The cyanuric chloride-modified scaffold undergoes two subsequent nucleophilic substitutions by R1- and R2-substituted nucleophiles, where R1 and R2 can be any selected functional groups. The first substitution occurs at 30°C for 24h and the second substitution at 80°C for 48h. Nucleophiles include amines, hydroxyls, thiols, etc, and careful selection of the R1 and R2 moieties is imperative to avoid cross-reactivity.

Multicomponent reactions are one-pot reactions which consist in the combination of three or more compounds to yield a single product. The resulting product retains all of the atoms of the starting materials, excepting the condensation products [14]. Both Petasis and Ugi reactions fall into this category.

Ugi reaction was reported in 1959 by Ivar Ugi and co-workers, and consists in the condensation between a primary or secondary amine, an aldehyde or ketone, a carboxylic acid and an isonitrile [16]. Ugi reaction on solid-phase is depicted in Figure 2.2.

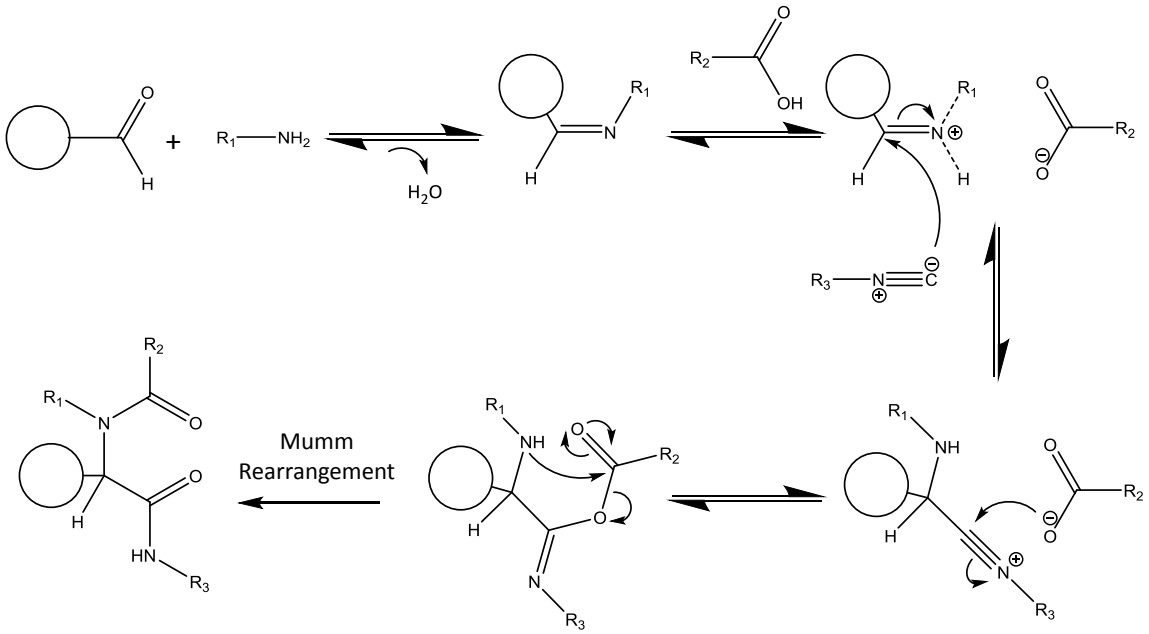


Figure 2.2 – Mechanism of Ugi reaction on solid-phase. The first step of the Ugi reaction consists in the condensation between the amine component and the aldehyde/ ketone (herein attached to the solid support) to form an imine, which in turn reacts with the carboxylic acid via an acid-base reaction to give an iminium ion. A nitrilium ion is then generated by the reaction between the iminium and the isocyanide component. The nitrilium ion further reacts with the carboxylate ion generated in the imine activation step, yielding an unstable imino-anhydride, which then undergoes a Mumm rearrangement (irreversible step) to generate the final Ugi product [17].

The Petasis borono-Mannich reaction consists of adding an aldehyde, an amine, and a boronic acid, being named after its discovery by Nicos A. Petasis in 1993 [18]. The reaction has been used for the synthesis of α-amino acids, amino alcohols, 2-hydroxymorpholines, 2H-chromenes, etc. Secondary amines are the most reactive, followed by bulky primary amines, but the employment of tertiary amines has been reported as well [19-21]. The presence of a hydroxyl group in the amine or aldehyde component has been recognized to be fundamental for the activation of the boronic acid. Therefore, aldehydes comprising hydroxyl or carboxyl groups, such as glycolaldehyde, glyoxylic acid and

salicylaldehyde are commonly reported [20, 22, 23]. Petasis and Zavialov have also successfully used α-keto acids, such as pyruvic acid, instead of the aldehyde component [21].

The reaction mechanism which is currently considered more energetically favourable is represented in Figure 2.3, but some mechanistic differences may be observed depending on the nature of each of the components [20, 22-24].

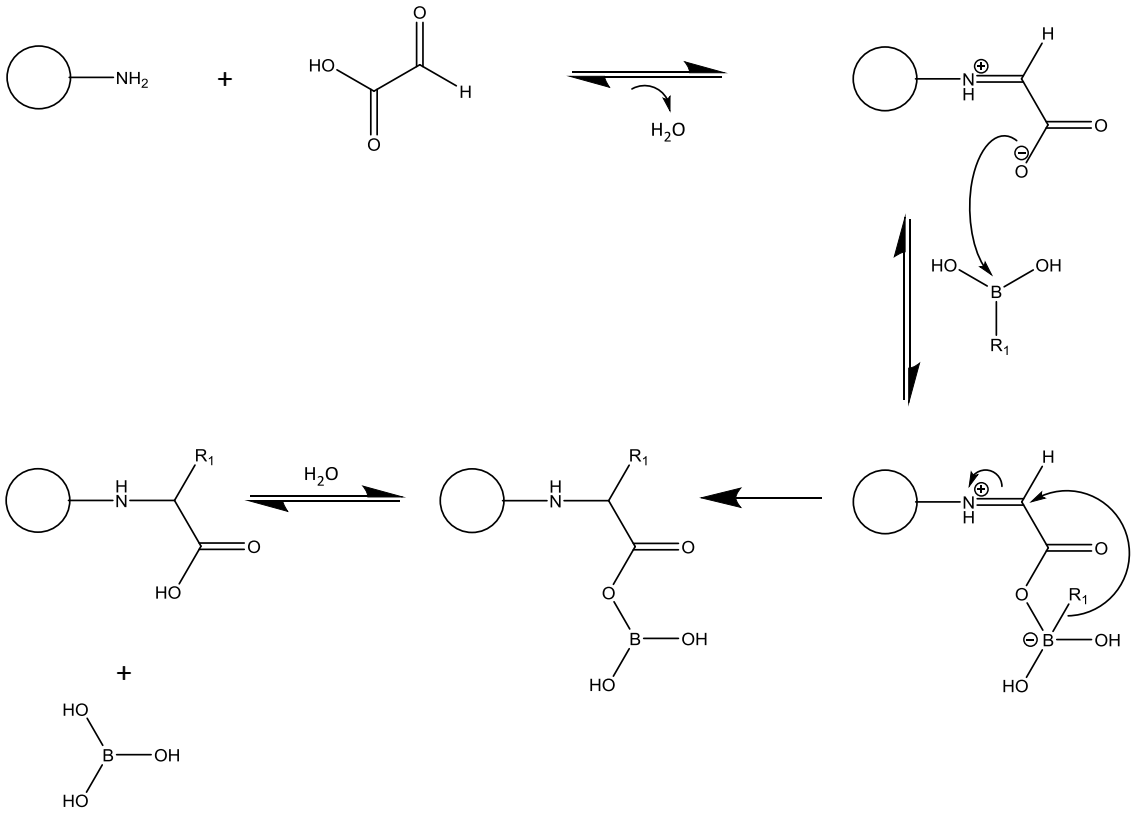


Figure 2.3 – Proposed mechanism of Petasis borono-Mannich reaction on solid-phase. In this example, the amine component is attached to the solid matrix and glyoxylic acid is used as the aldehyde component. The first step of the reaction consists of a condensation between the amine and the aldehyde components yielding a zwitterion species. The carboxylate of the zwitterion then reacts by nucleophilic addition with the boron from the boronic acid, forming a negatively charged tetra-coordinated boronate intermediate – the "ate complex". The boronic acid substituent (R_1) then migrates to the iminium's electrophilic carbon, which is the irreversible step of this reaction. Both the product and boric acid are then generated by a hydrolysis reaction.

As illustrated in Figure 2.3, in this case the Petasis product generates a free carboxylic acid (deriving from the glyoxylic acid component), which can be further employed as a component in other reactions, such as the Ugi reaction. The tandem Petasis-Ugi reaction tremendously increases molecular diversity, by allowing the incorporation of a higher number of functional groups.

2.2. Materials and Methods

2.2.1. Chemicals

All reagents were of the highest purity available and the solvents were pro-analysis.

1-Pyrenemethylamine hydrochloride (AP), 2-(ethylthio)ethylamine (A10), 2-ethylhexanal (B4), 3-(4-hydroxyphenyl)propionic acid (C3), 3-(ethylthio)propanoic acid (C5), 3-(methylthio)propionaldehyde (B5), 3-indoleacetic acid (C7), 3-thienylboronic acid, 4-imidazolecarboxaldehyde (B1), 4-imidazolecarboxylic acid (C1), agmatine sulfate salt (A7), amino-2-propanol (A2), ammonium bicarbonate, ammonium hydroxide solution (NH_4OH), bicinchoninic acid kit for protein determination (BCA), cadaverine dihydrochloride (A5), coumarin-3-carboxylic acid (C6), cyanuric chloride, epichlorohydrin, glyoxylic acid monohydrate, histamine (A8), indole-3-carboxaldehyde (B7), isopentylamine (A9), isopropyl isocyanide, *m*-xylylenediamine (A11), *N*,*N*-dimethyl-4,4'-azodianiline (A1), ninhydrin, phenethylamine (A3), phenol, phenylacetaldehyde (B3), phenylboronic acid, potassium cyanide, propionaldehyde (B2), pyrene-1-boronic acid, pyridine, sodium bicarbonate ($NaHCO_3$), sodium periodate, sodium thiosulfate, succinamic acid (C2), tetrahydrofuran-3-carboxaldehyde solution (B6), tris(hydroxymethyl)aminomethane (A12), tryptamine (A6), tyramine (A4), and γ-aminobutyric acid (C4) were acquired from Sigma-Aldrich. 2-Propanol and methanol (MeOH) were obtained from Roth, 4-aminomethylphenylboronic acid hydrochloride from Synthonix, and acetic acid glacial from Pronalab. Acetone and dimethylformamide (DMF) were purchased from VWR. Acetic anhydride, ethanol absolute PA, hydrochloric acid 37% (HCl), nitric acid, sodium acetate and sodium hydroxide (NaOH) were acquired from Panreac. The peptides Ser-Gln-Val-Phe-Pro-Trp (SW6) and pSer-Gln-Val-Phe-Pro-Trp (SW6-P) were > 98% pure and were purchased from Genecust and Caslo.

2.2.2. Chromatographic materials

Sepharose™ CL-6B was acquired from GE Healthcare. Captiva 96-well 20u filter plate with the respective duo seal 96-well Pl seal and Captiva 96-well plate cover were purchased from Agilent Technologies. Half-area UV-Star® 96-well microplates, black immunograde 96-well microplates, and 96-well transparent microplates were obtained from Greiner Bio-One, Brand, and Sarstedt, respectively.

2.2.3. Instrumentation

Ligand synthesis was carried out in an IKA KS 4000 ic control shaker. Inductively Coupled Plasma - Atomic Emission Spectroscopy (ICP-AES) analysis for the detection of boron and sulphur in ligand-functionalized agarose was conducted in a Horiba Jobin-Yvon Ultima Spectrophotometer at an impressed potential of 1200 kW. Fluorescence microscopy assays used an Olympus BX51 microscope (400X magnification), a U-MWB filter (λexc = 460-490nm; λem = 515-700nm), an UPlanFL objective, an U-RFL-T lamp, and Cell View software for monitoring. Fluorescence and absorbance readings in 96-well microplates were conducted in a Tecan F200 Microplate Reader using a λexc = 280 (20) nm – λem = 340 (35) nm filter, and 560 (10) nm filter, and a 280 (5) nm filter (Tecan).

2.2.4. Methods

2.2.4.1. Epoxy-activation of agarose

Sepharose™ CL-6B was washed with distilled deionised water (dd water) (10x resin volume) in a sinter funnel using vacuum suction, and then resuspended in dd water (1 mL/ g moist agarose) and NaOH 10 M (0.04 mL/ g moist agarose). The suspension was incubated for 30 min at 34°C with orbital shaking (230 rpm). Epichlorohydrin was then added in the proportion of 0.072 mL/ g moist agarose and the mixture was left for another 3 h at 34°C in the orbital shaker (230 rpm). In the end, the resin was washed with dd water (10x resin volume). In order to determine the amount of epoxy groups, 1 g of epoxy-activated agarose was incubated with 3 mL of an aqueous solution of sodium thiosulfate 1.3 M for 20 min at room temperature (RT) with agitation. Epoxy groups were then quantified by titration with 0.1 M HCl. Typical values were 21 ± 1 μmol epoxy/ g moist agarose.

2.2.4.2. Functionalization of agarose with 4-aminomethylphenylboronic acid

4-Aminomethylphenylboronic acid hydrochloride (3 equivalents (eq.) molar excess relative to epoxy) was dissolved in dd water with an eq. of NaOH, in order to make the amine reactive towards the epoxy groups on agarose. Then, the boronic acid solution was added to the epoxy-activated agarose (1 mL solution/ g moist agarose) and the reaction was left for 24h at 60°C with orbital agitation (200 rpm). The resin was thoroughly washed with dd water by vacuum filtration.

2.2.4.3. Petasis reaction on boronic acid modified agarose

Agarose functionalized with 4-aminomethylphenylboronic acid was washed with 25% (v/v) ethanol/ dd water (5x resin volume) and 50% (v/v) ethanol/ dd water (5x resin volume). A mixture of 1-pyrenemethylamine hydrochloride and glyoxylic acid monohydrate (5 molar eq. of each compound in excess relative to the epoxy) in 50% (v/v) ethanol/ dd water was added and the reaction occurred for 48h at 60°C with orbital shaking (1 mL/ g moist resin). NaOH was added to the amine component in the proportion of 1 molar eq. NaOH/ molar eq. of 1-pyrenemethylamine hydrochloride. 1-Pyrenemethylamine hydrochloride had been previously dissolved in a small amount of DMF. Ultimately, the functionalized resin was washed first with 50% (v/v) ethanol/ dd water (5x resin volume) and finally with dd water (10x resin volume).

2.2.4.4. Functionalization of agarose with amine groups

Epoxy-activated agarose was resuspended in 5 M NH_4OH (1.5 mL/ g moist resin) and incubated overnight in the orbital shaker at 40°C with agitation at 200 rpm. Aminated agarose was then washed with dd water (10x resin volume). The presence of amines was confirmed by the Kaiser test [25], which consists of adding 50 µl of 5% ninhydrin in ethanol (w/v), 50 µl of 80% phenol in ethanol (w/v) and 50 µl of 2% 0.001 M potassium cyanide in pyridine (v/v) to 0.1 g of moist aminated agarose and 0.9 mL dd water and heating the samples in a water-bath at 100°C for 5 min. The test is based on the reaction between ninhydrin and primary amines, yielding a purple/blue colour.

2.2.4.5. Petasis reaction on aminated agarose

Aminated agarose was washed with 25% (v/v) ethanol/ dd water (5x resin volume) and 50% (v/v) ethanol/ dd water (5x resin volume). Glyoxylic acid monohydrate and phenylboronic acid (5 molar eq. of each compound in excess relative to the epoxy) in 50% (v/v) ethanol/ dd water (0.5 mL of each/ g moist agarose) were subsequently added and the reaction occurred for 48h at 60°C with orbital shaking (200 rpm). In the end, the Petasis-scaffolded agarose was washed with 50% ethanol/dd water (v/v) (5x resin volume) and dd water (10x resin volume) by vacuum suction.

2.2.4.6. Blocking of unreacted amines on Petasis-functionalized agarose

Petasis agarose was washed in a sinter funnel with 20% to 100% DMF (in 20% increments). The resin was resuspended in a solution of 10% (v/v) acetic anhydride in DMF and incubated for 24h at room temperature with orbital agitation (200 rpm). In the end, the resin was washed with 100% to 20% DMF in 20% decrements and dd water (10x resin volume).

2.2.4.7. Petasis-Ugi library synthesis

After blocking the amines, Petasis-functionalized agarose was washed with MeOH (20% to 100% MeOH in 20% increments), resuspended in MeOH (1 mL/ g moist resin), and distributed for 84 wells (0.25 g resin/ well) of a 96-well filter plate. A 1 mL pipette tip was end cut in 4 mm and 0.5 mL of resin slurry was pipetted to each well, which corresponds to approximately 0.25 g resin. This step is critical to guarantee a uniform distribution of resin in the filter plate. The slurry was left to drain by gravity in order to remove the solvent. In the end, the 96-well filter plate was end capped with a proper seal. In the meantime, aldehyde and amine components (5 molar eq. of each relative to epoxy) were dissolved in MeOH (0.25 mL/ component/ well) and were reacted for 2h at 60°C in the orbital shaker (200 rpm), in order to facilitate the imine formation required for the Ugi reaction. After this period, each mixture of amine and aldehyde (0.5 mL/ well) was added to the Petasis resin, along with isopropyl isocyanide (5 molar eq. relative to epoxy). The filter plate was covered with a plate cover and the reaction was left for 48h at 60°C in the orbital shaker (100 rpm). The Petasis-Ugi ligands were then washed with 0.75 mL of 100% to 20% MeOH (in 20% decrements) and then with dd water (10x 0.75 mL). The amines used were A1, A2, A3, A4, A5, A6, A7, A8, A9, A10, A11, and A12. The aldehydes were B1, B2, B3, B4, B5, B6, and B7. A1, A7, and A12 were dissolved in 50% (v/v) DMF/ MeOH. NaOH 1M was added to A5 (1 molar eq.) and A7 (2 molar eq.)

2.2.4.8. Functionalization of agarose with aldehyde groups

Epoxy-activated agarose was resuspended in NaOH 5 M (1 mL/ g moist resin) and incubated overnight at 40°C with agitation at 200 rpm. The resin was then thoroughly washed with dd water using vacuum filtration and resuspended in sodium periodate 0.1 M (1 mL/ g moist resin), followed by incubation at 45°C and 200 rpm for 6h. After that, the resin was washed with dd water (10x resin volume).

2.2.4.9. Ugi library synthesis

The procedure for Ugi library synthesis is similar to the one described for Petasis-Ugi library (§2.2.4.7), but in this case aldehyde-functionalized agarose was distributed by the 96-well filter plate (0.25 g/ well). Amines A1 to A12 (5 molar eq. relative to epoxy, 0.25 mL/ well) were added to each column of the filter plate; the plate was sealed and incubated for 2h at 60°C in the orbital shaker (100 rpm). After that period, the carboxylic acid components (5 molar eq. relative to epoxy, 0.25 mL/ well) were added to each row of the filter plate, and isopropyl isocyanide (5 molar eq. relative to epoxy) was added to the 84 wells. The up and end-capped filter plate was then incubated for 48h at 60°C in the orbital shaker (100 rpm). The rest of the procedure is identical to §2.2.4.7. The carboxylic acids used were C1, C2, C3, C4, C5, C6, and C7. C1, C4 and C6 were dissolved in 50% (v/v) DMF/ MeOH, while the rest of the carboxylic acids were dissolved in MeOH.

2.2.4.10. Triazine library synthesis

Aminated agarose was washed with and resuspended in a cold solution of 50% (v/v) acetone/ dd water with 1 molar eq. of $NaHCO_3$ relative to epoxy (1 mL/ g moist agarose). Cyanuric chloride (5 eq. molar excess relative to epoxy content) was dissolved in acetone (8.6 mL/ g cyanuric chloride) and added to the aminated resin, followed by incubation for 1h in ice at 200 rpm. After incubation, the triazine-functionalized resin was washed with acetone (2x resin volume), 50% (v/v) acetone/dd water (3x resin volume) and dd water (5x resin volume). The cyanuric chloride-functionalized agarose was then distributed by 64 wells of a 96-well filter plate (0.25 g/ well), and the solvent was left to drain by gravity. After end-capping the filter plate, amines A1 to A8 were added to each column of the filter plate (2 molar eq. of each relative to epoxy; 0.5 mL/ well), an upper cover was placed and the filter plate was incubated for 24h at 30°C with agitation (100 rpm). Afterwards, the ligands were washed with the solvent in which each amine was dissolved (5x 0.75 mL) and dd water (5x 0.75 mL), and in the end the solvent was left to drain by gravity. The filter plate was end-capped, and amines A1 to A8 were then added to each row of the filter plate (5 molar eq. of each relative to epoxy; 0.5 mL/ well). The filter plate was sealed and incubated for 48h at 80°C in the orbital shaker (100 rpm). In the end, the ligands were washed with the solvent in which each amine was solubilized (5x 0.75 mL), 0.1 M hydrochloric acid (HCl) (1x 0.75 mL), dd water (1x 0.75 mL), 0.1 M NaOH in 30% isopropanol (v/v) (regeneration buffer; 0.75 mL), and finally dd water (5x 0.75 mL). All the amines were dissolved in dd water, with the exception of A1, A6 and A7, which were dissolved in 50% (v/v) DMF/ dd water.

$NaHCO_3$ 1M (1 molar eq. relative to epoxy) was added to each amine. NaOH 1M was added to A5 (1 molar eq. relative to A5) and A7 (2 molar eq. relative to A7).

2.2.4.11. Sample preparation for ICP-AES

Petasis-Ugi, Ugi and triazine ligands were prepared accordingly to the methodologies described above, but all reactions were carried out in flasks instead of 96-well filter blocks. The components used for the synthesis of Petasis ligands were: (i) aminated agarose, 3-thienylboronic acid, and glyoxylic acid monohydrate for ligand D3; and (ii) 4-aminomethylphenylboronic acid-functionalized agarose, 1-pyrenemethylamine hydrochloride and glyoxylic acid monohydrate for ligand D7. The components used for the synthesis of Petasis-Ugi ligands were: D3-functionalized agarose, 2-(ethylthio)ethylamine, propionaldehyde, and isopropyl isocyanide for ligands D3A10B2 (without the intermediate step of blocking unreacted amines) and D3bA10B2 (after acetylating free amines on D3-functionalized agarose). The components used for the synthesis of Ugi ligand A10C2 were aldehyde-functionalized agarose, 2-(ethylthio)ethylamine, succinamic acid and isopropyl isocyanide. A symmetric triazine ligand was synthesized using 2-(ethylthio)ethylamine (ligand A10A10). Ligand-functionalized agarose was dried at 100°C for 5 days. Agarose was weighted before and after the drying process, in order to determine the weight difference between moist and dried agarose. A known amount of dried agarose was weighted and put in a glass flask to which 1 mL of 1:3 (v/v) nitric acid/ HCl was added. The flask was heated in a water-bath at 60°C until the agarose was solubilized, and then 2 mL of dd water was added to the sample. Samples were taken for ICP-AES analysis.

2.2.4.12. Sample preparation for fluorescence microscopy

Petasis-Ugi, Ugi and triazine ligands were synthetized in flasks. The components used for the synthesis of Petasis ligand D4 were aminated agarose, pyrene-1-boronic acid, and glyoxylic acid monohydrate. The components used for the synthesis of Petasis-Ugi ligands D4APB2 and D4bAPB2 were D4-functionalized agarose, 1-pyrenemethylamine hydrochloride, propionaldehyde, and isopropyl isocyanide. The components used for the synthesis of Ugi ligand APC2 were aldehyde-functionalized agarose, 1-pyrenemethylamine hydrochloride, succinamic acid and isopropyl isocyanide. Triazine ligand APAP was synthesized using 1-pyrenemethylamine hydrochloride as the amine nucleophile. Samples were pipetted to a microscope slide analysed by fluorescence microscopy in an Olympus BX51 microscope.

2.2.4.13. Screening of solid-phase combinatorial libraries

Petasis-Ugi, Ugi and triazine libraries were washed with regeneration buffer followed by dd water (3 cycles of washes, 0.75 mL/ well). Library equilibration was performed by washing the libraries extensively with the binding buffer (8x 0.75 mL/ well on average), until the UV absorbance of the washes at 280 nm was below 0.005 (last washes were collected in a half-area UV-Star® 96-Well microplate). Two binding buffers were tested: 50 mM acetate pH 4 and 50 mM ammonium bicarbonate pH 8. SW6 and SW6-P peptides were solubilized in binding buffer (1 mg/ mL) and 0.25 mL of each peptide was loaded per well. The ligand-functionalized resins were incubated with the peptides for 1h at room temperature with manual shaking. The libraries were then washed with binding buffer (8x 0.25 mL). Both the flow-through and washes were collected in 96-well black microplates by centrifugation at 600 rpm for 20 s, and the fluorescence intensity was read on a microplate reader using a λexc = 280 (20) nm – λem = 340 (35) nm filter. After that, 25 µl of each sample was transferred from the black microplate to a 96-well transparent microplate to perform the BCA assay. The assay consists of adding 200 µl of BCA working reagent to 25 µl of sample and incubating the mixture 30 min at 37°C in the dark. The absorbance of the samples was then determined in the microplate reader at 560 nm.

2.3. Results and Discussion

Three combinatorial libraries of synthetic biomimetic ligands were synthesized in solid-phase using three different chemical scaffolds and agarose as the solid support. These libraries were designed based on the study of key structural elements of different phosphoprotein-binding domains (PBDs) and resulted in 232 small molecules. The ligands were further screened against a pSer-Gln-Val-Phe-Pro-Trp peptide (named SW6-P), which contains the consensus amino acid sequence - pSer-X-X-Phe (X is any amino acid) - known to specifically bind the Brca1 C-terminal (BRCT) domain.

2.3.1. Synthesis of Triazine, Ugi and Petasis-Ugi based scaffolds

The synthesis of triazine and Ugi-based ligands on agarose beads has been previously reported, and their experimental conditions optimized [15, 26]. Although the tandem Petasis-Ugi reaction is known since 2003, up till now it was never attempted to use this scaffold for proteomic experiments, neither was its synthesis on chromatographic supports. Portlock and co-workers used this reaction for the solid-phase synthesis of small molecular weight compounds for drug discovery using RINK amine and RINK isonitrile resins [27]. Besides having a different application, they have used a different methodology. In their approach, the solid support is introduced as one of the Ugi components, while here it is used as one of the Petasis components.

The feasibility of Petasis-Ugi reaction on agarose was assessed by using both boronic acid- and amine-functionalized agarose. If Petasis reaction goes to completion, then the boron element would not be present in the final product, since only its R-substituent (R-B$(OH)_2$) is incorporated upon the reaction. Agarose functionalized with 4-(aminomethyl)phenylboronic acid was reacted with glyoxylic acid and 1-pyrenemethylamine, and the amount of boron in Petasis-scaffolded agarose (mol B/ g agarose) was quantified by ICP-AES. These three components were selected due to the absence of problematic functional groups that could potentially interfere with the accurate course of the reaction: glyoxylic acid is the simplest α-keto acid and the other two components only possess aromatic substituents. ICP results show that only 30% of the 4-(aminomethyl)phenylboronic acid actually reacts with the agarose, and 70% of this reacted to yield the Petasis product (Figure 2.4). The 30% of the boronic acid that did not react is sufficient to hamper the scaffold applicability in chromatographic processes, since the solid support should be as inert as possible. Boronic acids are strong Lewis acids and easily react in acid-base reactions. They are able to form covalent, ionic and hydrogen bonds, and are known to coordinate sugars and amino acids comprising 1,2- or 1,3- Lewis

base donors, such as hydroxyl, carboxyl or imidazole groups [28]. Unreacted boronic acids could establish non-specific bonds with the target proteins, hampering the applicability of the affinity support. In light of this, a different approach using aminated agarose was then pursued.

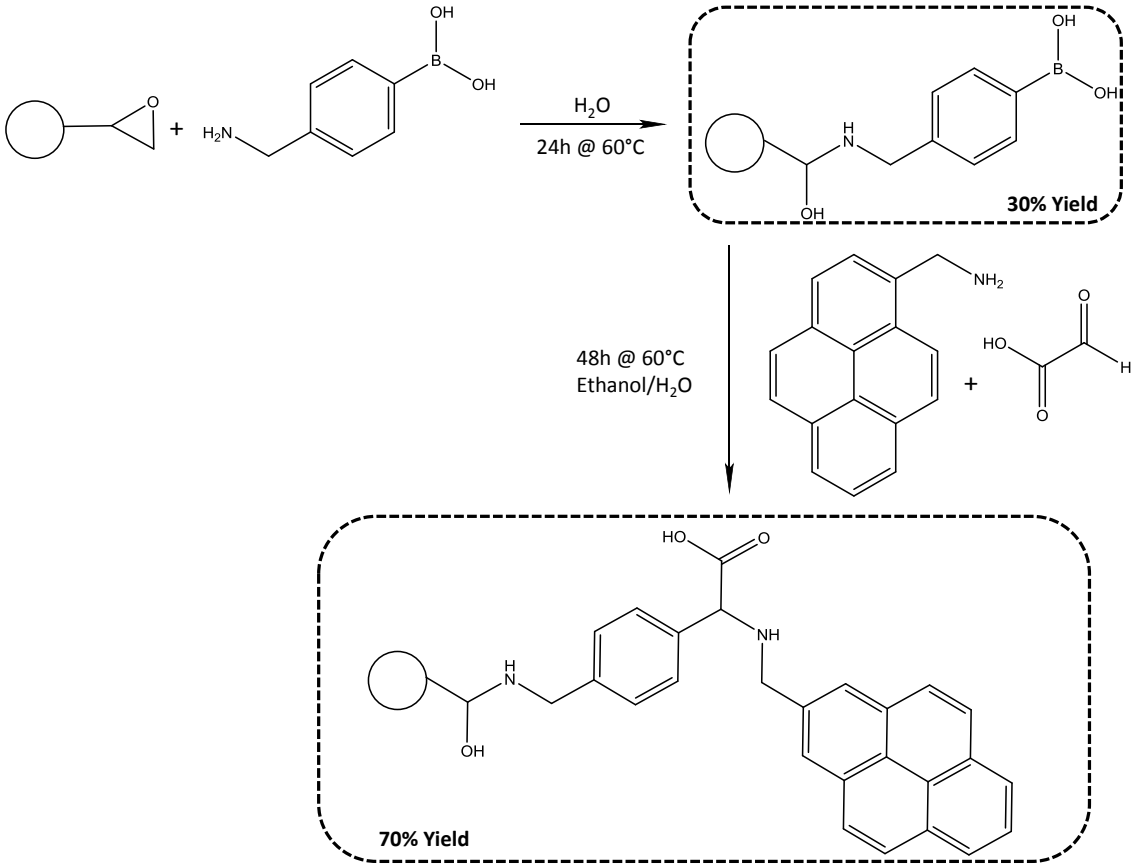


Figure 2.4 – Petasis reaction on boronic acid functionalized agarose. Epoxy-activated agarose was functionalized with 4-(aminomethyl)phenylboronic acid. The boronic acid-modified agarose was then used as one of the components in Petasis reaction. The amine and aldehyde components were 1-pyrenemethylamine and glyoxylic acid, respectively. Yields of reaction were calculated by ICP-AES based on the amount of boron present in the solid support (mol B/ g support).

Aminated agarose was reacted with 3-thienylboronic acid and glyoxylic acid, and both boron and sulphur were quantified in the Petasis-scaffolded agarose by ICP-AES (Figure 2.5). The yield of the reaction is determined by the incorporation of sulphur in the solid support (mol S/ g moist agarose), and corresponds to 18% relative to the epoxy groups on agarose[4]. No boron was detected in the Petasis-scaffolded agarose, which indicates that the boronic acid is not binding in a non-specific manner to the solid support, and that Petasis reaction is plausibly proceeding as expected.

[4] The determination of amine groups on agarose by the Kaiser test (which is a qualitative test) is not as reliable as the determination of epoxy groups by titration with HCl. Therefore, as aminated agarose derives from epoxy-activated agarose, all calculations were performed relative to epoxy-activated agarose.

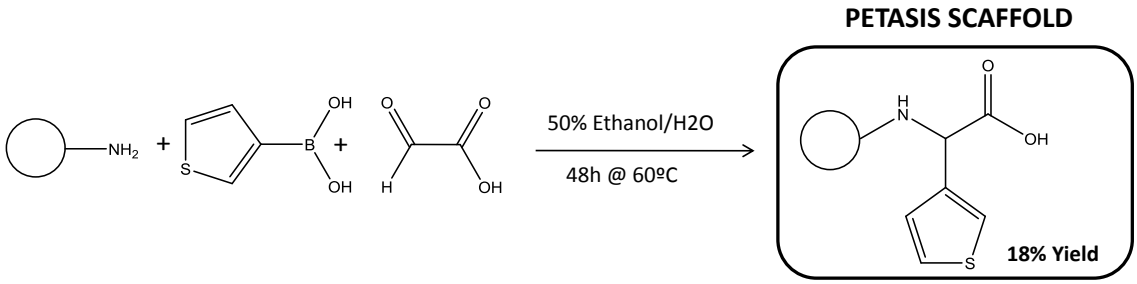


Figure 2.5 – Petasis reaction using aminated agarose as solid support. Approximately 4 µmol of sulphur were incorporated per g of moist agarose, which corresponds to a reaction yield of 18%. No boron was detected in the final product, which is consistent with the occurrence of the Petasis reaction. These results were obtained by ICP-AES.

The extent of multicomponent reactions, as the Petasis reaction, is highly dependent on the structure and electronic properties of each of the components and on the reaction conditions, such as the temperature and solvent selected. Petasis reaction has been performed in a variety of solvents, such as ethanol, dichloromethane, MeOH, toluene, DMF, etc [20]. Régnier and co-workers obtained good results by using microwave irradiation in a solvent-free system as well [29]. A low cost and environmentally friendly alternative is the use of water, which has been reported to be a suitable media for this reaction. The reaction kinetics may be significantly influenced by the use of this solvent, since it presents large surface tension, high specific heat capacity, high polarity, high cohesive energy, and the ability to form hydrogen-bonds [22, 23].

In the present work a mixture of 50% (v/v) ethanol/ water was selected for the Petasis reaction, since the use of protic solvents is reported to increase the reactivity of primary amines [30]. It is difficult to select a universal solvent system and the appropriate temperature to perform these reactions, since the overall goal of this work is to build universal scaffolds, which can then be tuned to target specific substrates by changing the individual components.

ICP-AES results demonstrate that there is a high amount of amines in the solid-support that did not react to yield the Petasis product, and these amines may compete with the free amines in solution in the subsequent Ugi reaction. In a library of synthetic ligands, the yields of reaction may be completely different from one ligand to another; therefore, it is fundamental to block the unreacted amines before performing the Ugi reaction on Petasis-scaffolded agarose (Figure 2.6).

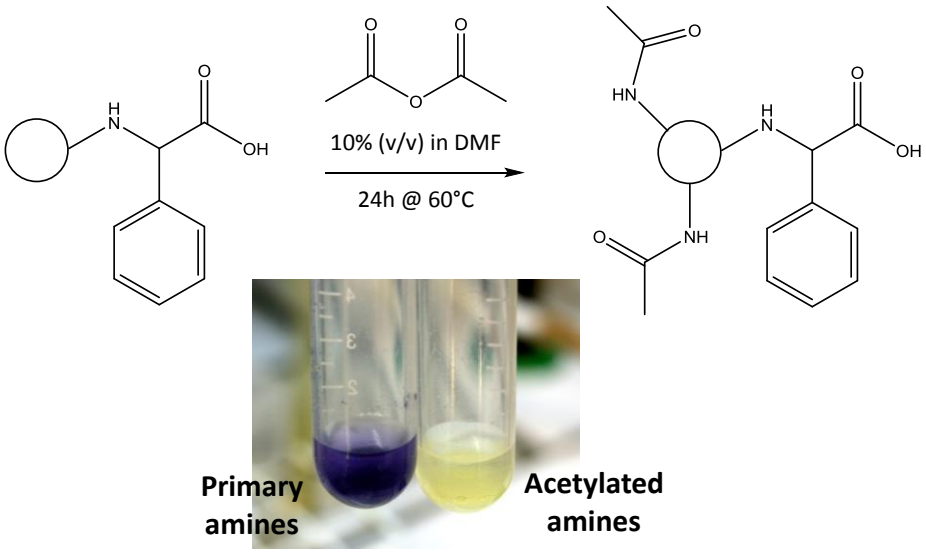


Figure 2.6 – Blocking primary amines with acetic anhydride. Unreacted amines on Petasis-scaffolded agarose were blocked using a 10% (v/v) solution of acetic anhydride in DMF. The success of the reaction was confirmed using Kaiser test. Petasis-scaffolded beads presented a purple colour characteristic of primary amines, while a whitish colour was observed after the blocking reaction. The Petasis reaction preceding the blocking reaction was performed using aminated agarose, phenylboronic acid and glyoxylic acid.

Aldehyde-functionalized beads could also be employed as the solid-phase component of the Petasis reaction, while using the amine and boronic acid in solution-phase. However, care must be taken upon the selection of the aldehyde to be immobilized, since it can be a determining factor for the success or failure of the reaction. Schlienger and co-workers performed the Petasis reaction on solid-phase using Wang polystyrene resins and observed that the most reactive aldehydes bear a α- or β-hydroxyl group [31]. This might also be explained by the mechanism of Petasis reaction highlighted in Figure 2.3.

The triazine, Ugi and Petasis-Ugi scaffolded agarose beads used in this work are illustrated in Figure 2.7.

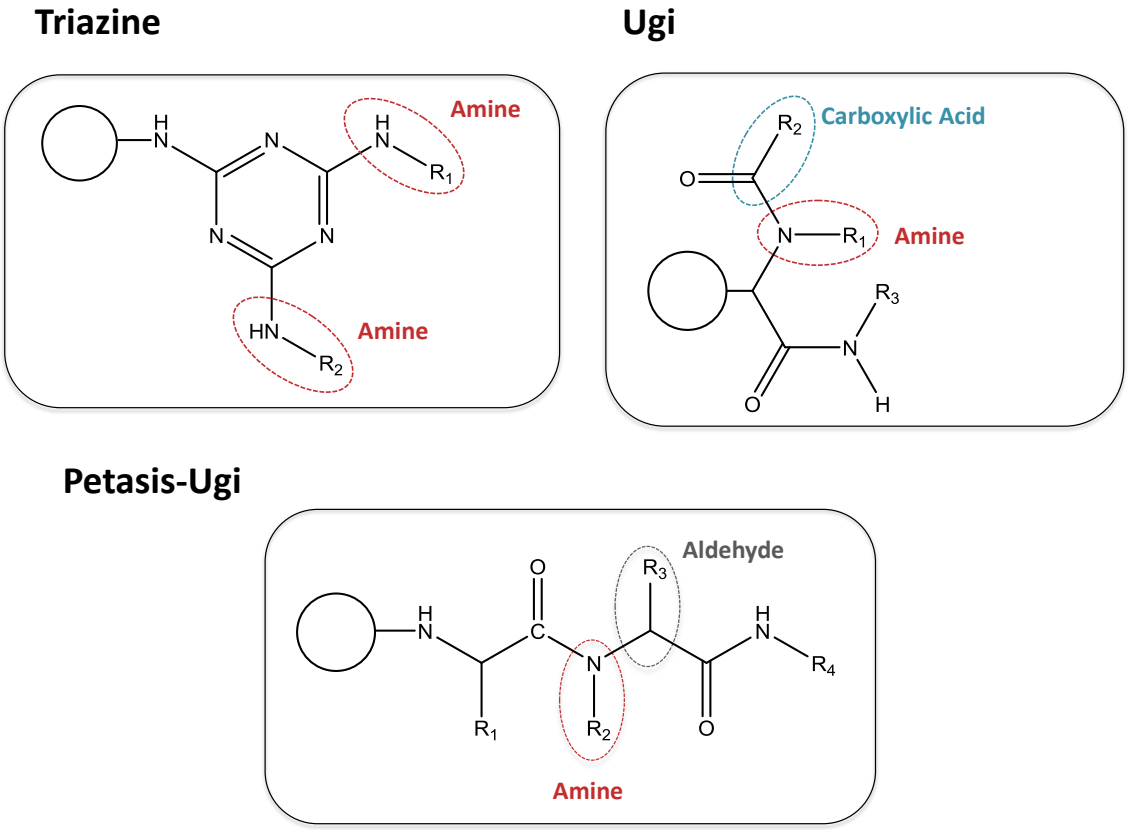


Figure 2.7 – Structures of triazine, Ugi and Petasis-Ugi based scaffolds. Triazine-scaffolded beads have two amine substituents (R1 and R2), which result from two sequential nucleophilic substitutions. Ugi scaffolds have three substituents derived from an amine (R1), a carboxylic acid (R2), and an isocyanide (R3). Petasis-Ugi scaffolds have four substituents derived from a boronic acid (R1), an amine (R2), an aldehyde (R3) and an isocyanide (R4). The components which were varied in this work are highlighted in colours; the others were kept constant.

As for Petasis-scaffolded beads, the success of triazine and Ugi reactions was also followed by the incorporation sulphur onto their structures, and subsequent detection by ICP-AES. The scaffolded beads used for these studies are depicted in Figure 2.8.

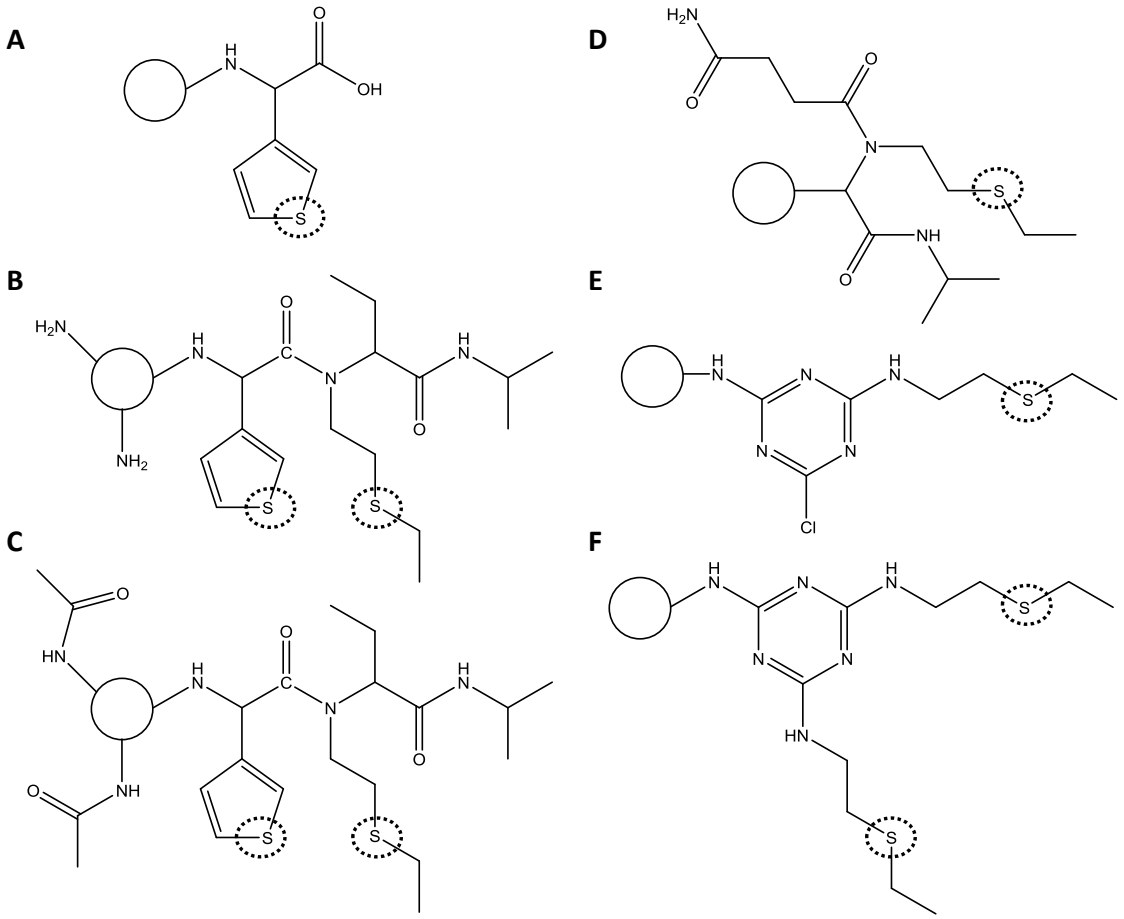


Figure 2.8 – Structures of scaffolded agarose beads used for ICP-AES studies. A) Petasis-scaffolded bead (D3). B) Petasis-Ugi-scaffolded bead with unreacted amines (D3A10B2). C) Petasis-Ugi-scaffolded bead after blocking free amines (D3bA10B2). D) Ugi-scaffolded bead (A10C2). E) Triazine-scaffolded bead after the first nucleophilic substitution (A10). F) Triazine-scaffolded bead after the second nucleophilic substitution (A10A10). The sulphur atom is highlighted by a dashed circle.

The results show that all the reactions are successfully performed on agarose (Figure 2.9). Petasis reaction (D3) presents the lower yield (18% relative to epoxy-activated agarose), but the subsequent Ugi reaction on Petasis-scaffolded beads is highly efficient. In fact, the Petasis-Ugi scaffolds present reaction yields between 90% (D3A10B2) and 100% (D3bA10B2) relative to the Petasis-scaffolded beads. The free amines do not seem to be interfering with the Ugi reaction, although it is possible to see a slight increase in the reaction yield for the support containing the acetylated amines (D3bA10B2). Ugi reaction on aldehyde-functionalized beads (A10C2) occurs with 94% yield, which corresponds to 19.8±0.8 µmol S/ g moist agarose. The first nucleophilic substitution on triazine ring (A10) allows the incorporation of 11.2±0.7 µmol S/ g moist agarose, which corresponds to a reaction yield of 53% relative to epoxy-activated agarose. Ligand A10A10 possess 21.6±1.6 µmol S/ g moist agarose, which means the second nucleophilic substitution occurs with a reaction yield of 100% relative to the first substitution. These results are in accordance with the fact that the energy

necessary for the second nucleophilic substitution is much larger than for the first, and elucidates why it is important to perform the first nucleophilic substitution at lower temperature (30°C) and the second nucleophilic substitution at higher temperature (80°C) in order to accurately control the reaction.

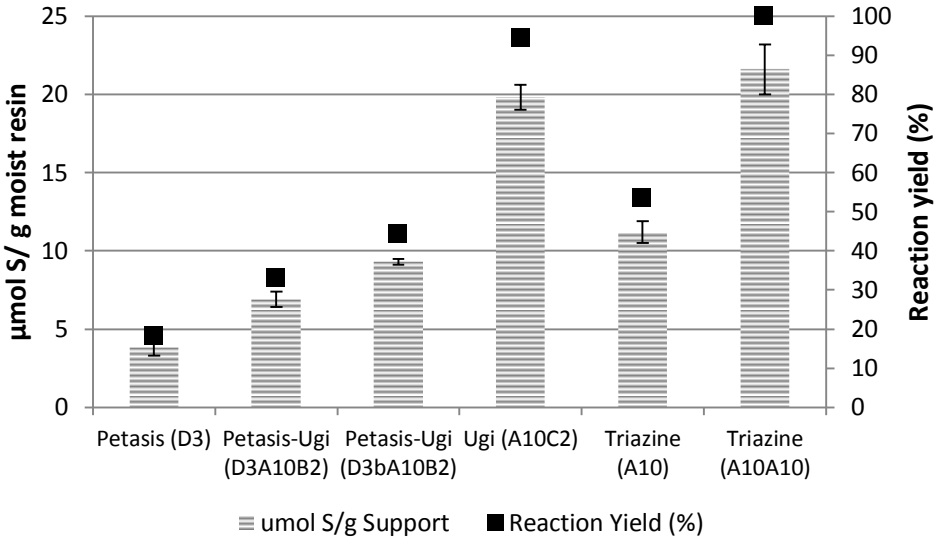


Figure 2.9 – Incorporation of sulphur onto scaffolded agarose beads. The reaction yields are relative to the moles of epoxy in epoxy-activated agarose. Petasis-scaffolded beads present the lower yield (18%). Petasis-Ugi scaffolds (D310B2 and D3bA10B2) present yields on 33% and 44% relative to epoxy-activated agarose, but the extension of the reaction relative to the Petasis-scaffolded beads is of 90% and 100%, respectively. Yields of 94% and 100% are observed for Ugi (A10C2) and triazine (A10A10) reactions on agarose beads, respectively.

On-bead scaffold functionalization was also evaluated by fluorescence microscopy by using a pyrene moiety as a reporter molecule (Figure 2.10). Pyrene displays a very particular feature, as it is able to form an excited-state dimer (excimer) when two pyrene molecules are superimposed in a proper position, increasing the maximum emission wavelength [32].

Fluorescence microscopy images are in accordance with ICP-AES results (Figure 2.11). All scaffolds present green light emission due to the presence of pyrene in their structures. Petasis-scaffolded beads present the lower fluorescence intensity, which might be explained by a low reaction yield of Petasis reaction, as previously observed. It is possible to see an increase in the fluorescence intensity from Petasis (D4) to Petasis-Ugi (D4APB2 and D4bAPB2) scaffolds, due to the presence of one and two pyrene molecules, respectively. However, it should be noted that fluorescence microscopy gives qualitative and not quantitative information. Fluorescence emission may be influenced by quenching, a phenomenon which consists in the reduction of fluorescence intensity. This may happen due to the interaction between the fluorophore and a quencher molecule, such as oxygen, halogens, amines,

etc, which can diffuse in solution or be present in a molecule in close proximity to the fluorophore; although, non-molecular mechanisms may also be involved in the quenching process [33]. This might explain why the fluorescence emitted by triazine scaffolds after the first nucleophilic substitution (AP) is much lower than the fluorescence observed after the second nucleophilic substitution (APAP), since there is a chloride atom in the vicinity of the chromophore in ligand AP. Another possible explanation can be related with the fact that the energy provided at 30°C is not enough for an efficient first nucleophilic substitution using 1-pyrenemethylamine in the triazine reaction. Amines, as any other functional compounds, present different reactivities. A high fluorescence emission is observed for both Ugi (APC2) and triazine ligand APAP, which is also in agreement with the ICP-AES results, where similar reaction yields were determined.

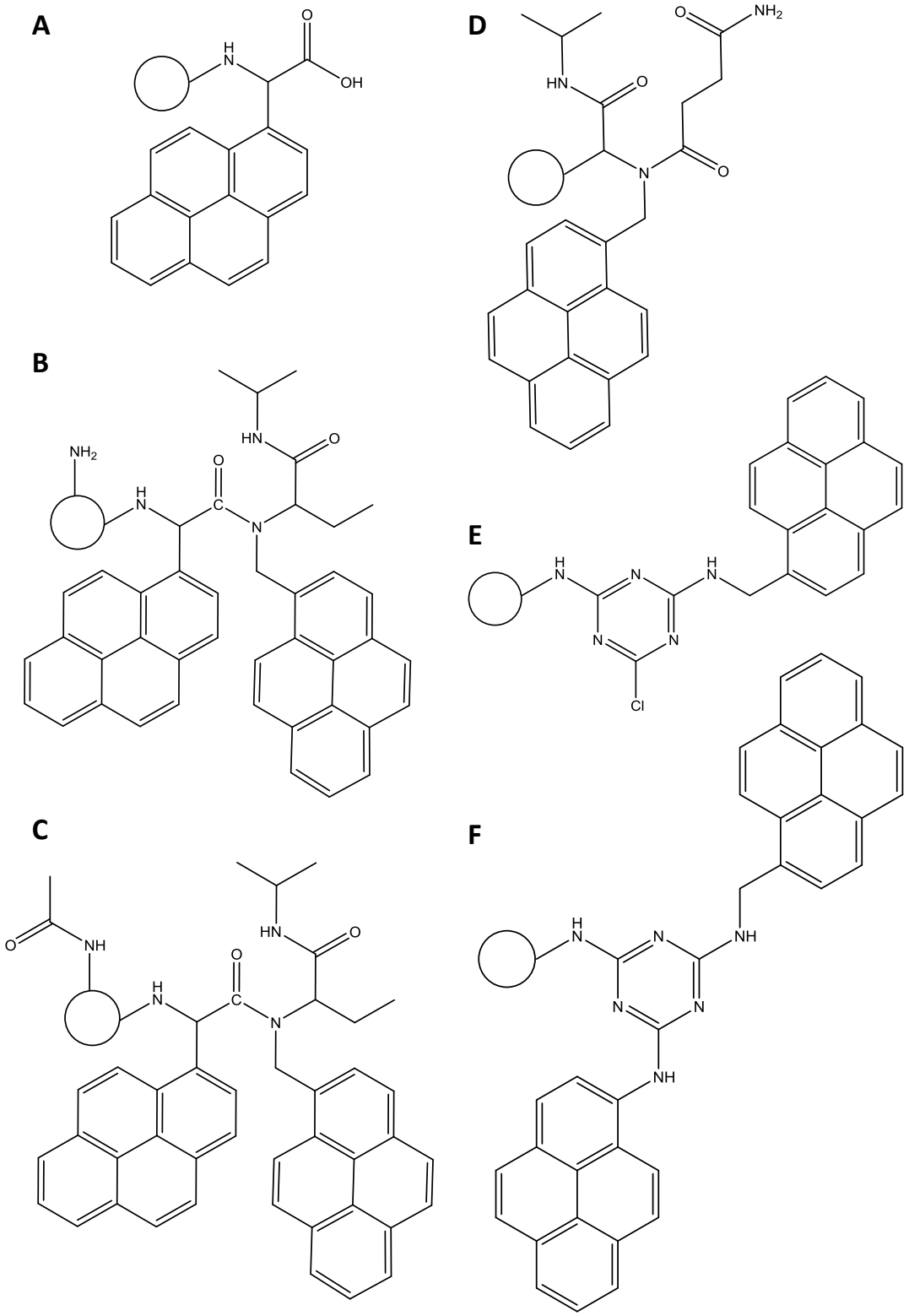


Figure 2.10 – Structures of scaffolded agarose beads used for fluorescence microscopy studies. A) Petasis-scaffolded bead (D4). B) Petasis-Ugi-scaffolded bead with unreacted amines (D4APB2). C) Petasis-Ugi-scaffolded bead after blocking free amines (D4bAPB2). D) Ugi-scaffolded bead (APC2). E) Triazine-scaffolded bead after the first nucleophilic substitution (AP). F) Triazine-scaffolded bead after the second nucleophilic substitution (APAP).

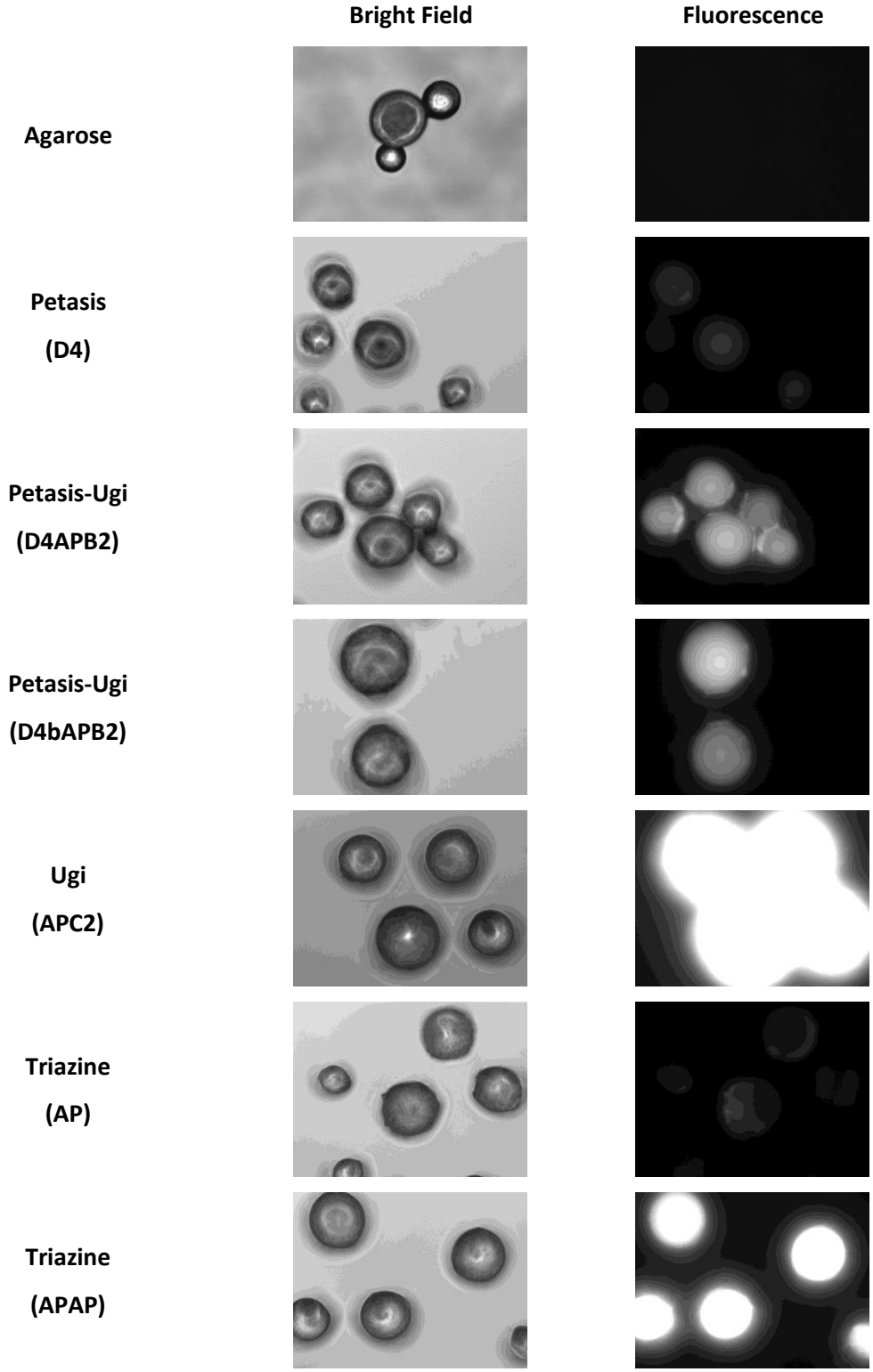


Figure 2.11 – Fluorescence microscopy results of scaffolded agarose beads functionalized with a fluorophore – pyrene. Agarose was used as negative control. The images were acquired using an Olympus BX51 microscope (400X magnification) with a U-MWB filter (λexc 460–490 nm; λexc 515–700 nm) and an Olympus U-RFL-T lamp. All images were acquired with an exposition time of 100 ms. The monitoring software was Cell-View System.

2.3.2. Library Design: Selection of amine, aldehyde, and carboxylic acid components for ligand synthesis

Structural studies of the ten human PBDs (§1.2) revealed the existence of prevalent amino acids which are directly involved in phosphate coordination: Arg, Lys, Tyr, Ser, His, Gly, and Asn (Figure 2.12).

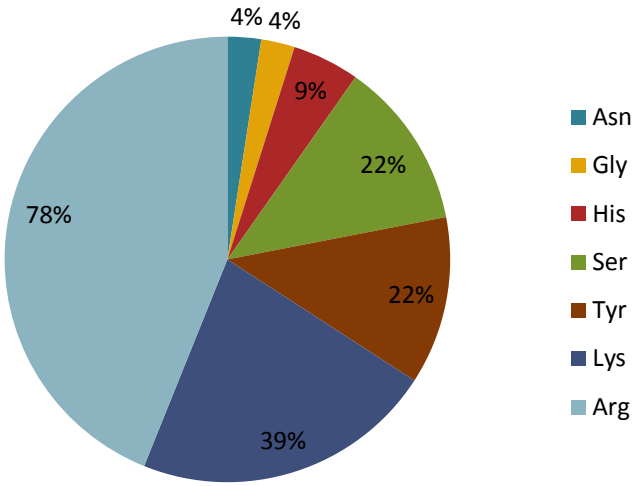


Figure 2.12 – Frequency of amino acids involved in the coordination of phosphate groups in PDBs. The amino acid frequencies were calculated based on the total number of amino acids establishing direct bonds to the phosphate moiety in PDBs. Arg and Lys clearly play key roles in phosphate binding.

Based on these results, different amine, aldehyde and carboxylic acid components were selected for the synthesis of the libraries. Some of these components were chosen to mimic the prevalent identified amino acids, while others were selected in order to increase library diversity. The main functional groups used in the libraries are amines and hydroxyl groups, but aliphatic and aromatic moieties were also introduced to mimic hydrophobic interactions.

Although the multicomponent reactions used herein allow the incorporation of three or more components, in this work we have started to vary only two components on each time per scaffold.

All components of Petasis reaction were kept constant. The reaction was performed using aminated agarose, phenylboronic acid and glyoxylic acid. As referred previously, the carboxylic acid resultant from this reaction will be used as one of the components on the subsequent Ugi reaction, where the isocyanide component was also kept constant. Isopropyl isocyanide was used due to its small size (preventing steric hindrance effects) and the fact that its R-substituent is a simple alkyl group

(preventing cross-reactivity). Therefore, in Petasis-Ugi scaffolds the variable components were the amine and the aldehyde (Table 2.1).

The Ugi reaction was performed on aldehyde-functionalized beads and again the isocyanide moiety was maintained constant. The components varied in this case were the amine and the carboxylic acid (Table 2.2).

Eight amines were used in the nucleophilic substitutions of triazine reaction, resulting in a 64 ligand library (

Table **2.3**). The triazine library is theoretically symmetric, because it is indifferent to use a certain amine in the first or second nucleophilic substitution. However, as amines have different reactivities, it is not guaranteed that there is 50% incorporation in the first nucleophilic substitution. Highly reactive amines may be incorporated in a higher extent at 30°C, while low reactive amines may present a much lower reaction yield.

Table 2.1 – Amine and aldehyde components of Petasis-Ugi library. Twelve amines (A1 to A12) and seven aldehydes (B1 to B7) were used to build a 84 ligand library. The compounds that were selected to mimic particular amino acid residues are next to the respective three letter amino acid code.

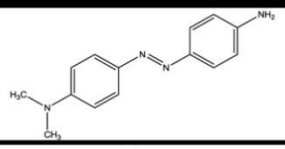
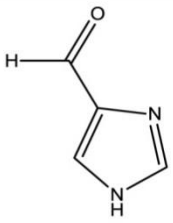
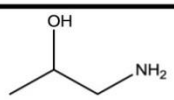
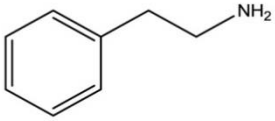
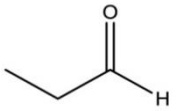
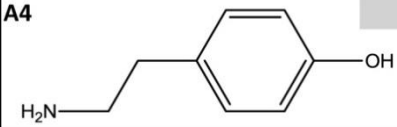
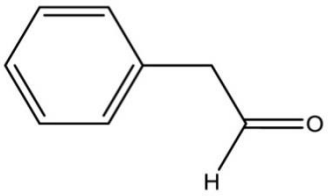
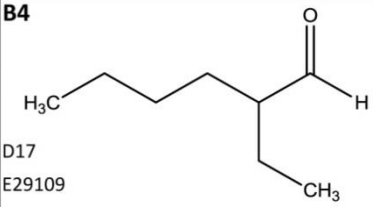
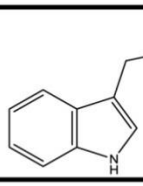
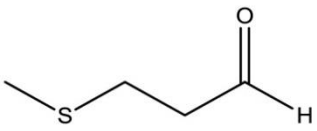
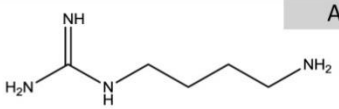
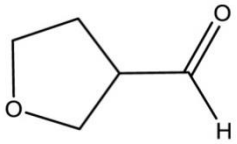
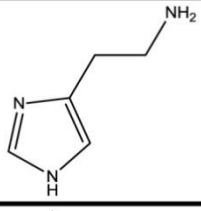
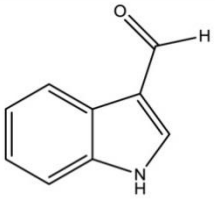
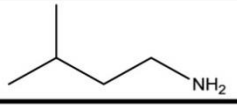
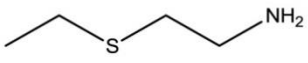
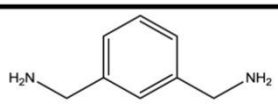
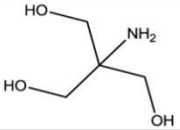
Amine		Aldehyde	
A1		B1	His 
A2	Ser 		
A3	Phe 	B2	
A4	Tyr 	B3	Phe 
A5	Lys 	B4	 D17 E29109
A6	Trp 	B5	Met 
A7	Arg 	B6	
A8	His 	B7	Trp 
A9	Leu 		
A10	Met 		
A11			
A12			

Table 2.2 - Amine and carboxylic acid components of Ugi library. Twelve amines (A1 to A12) and seven carboxylic acids (C1 to C7) were used to build a 84 ligand library. The compounds that were selected to mimic particular amino acid residues are next to the respective three letter amino acid code.

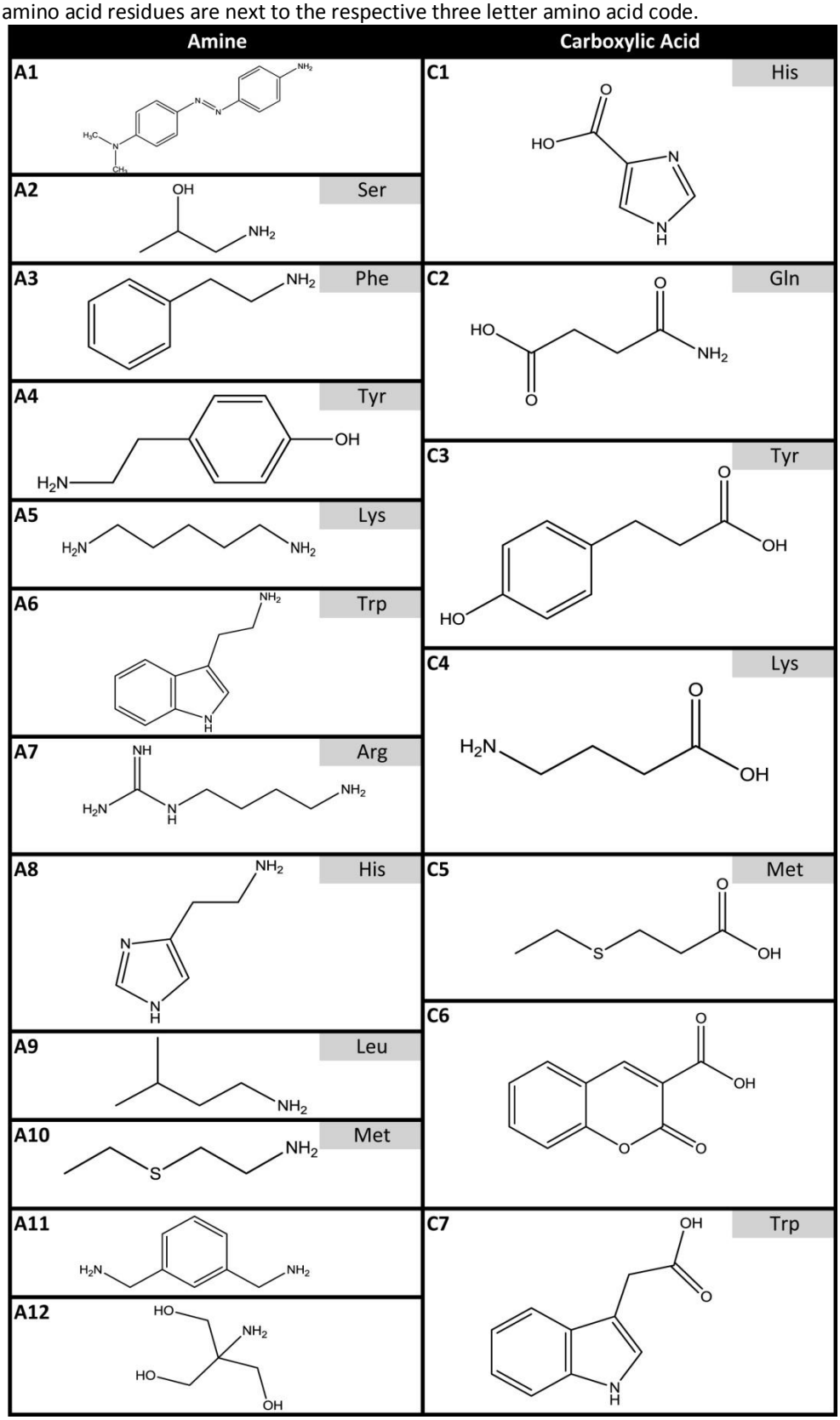
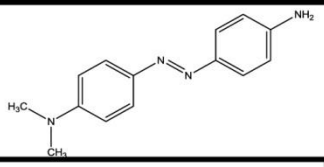
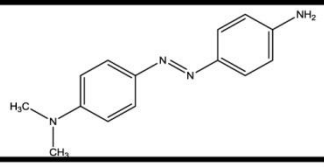


Table 2.3 - Amine components of triazine library. Eight amines were used to build an 84 ligand library. The amines used in the first and second nucleophilic substitutions are represented in the first and second columns, respectively. The compounds that were selected to mimic particular amino acid residues are next to the respective three letter amino acid code.

Amine		Amine	
A1		A1	
A2	Ser	A2	Ser
A3	Phe	A3	Phe
A4	Tyr	A4	Tyr
A5	Lys	A5	Lys
A6	Trp	A6	Trp
A7	Arg	A7	Arg
A8	His	A8	His

The five most prevalent amino acids (Arg, Lys, Tyr, Ser, His) are represented in all three libraries.

2.3.3. Screening of solid-phase combinatorial libraries

All libraries were screened against the SW6-P peptide and its non-phosphorylated version SW6. As the optimal binding condition is unknown, screenings were performed at pH 4 and pH 8, due to the fact that pSer (pKa_1=2.19 (PO_3H_2), pKa_2=5.78 (PO_3H^-), pKa_3=9.85 (NH_3^+)) is known to have one negative charge on the phosphate group at pH 4 and two negative charges in the phosphate group at pH 8 [34]. The non-bound peptide was quantified in a 96-well microplate reader by two methods: Trp fluorescence emission and the BCA assay. Among the aromatic amino acids, Trp is the one displaying the largest extinction coefficient, with absorption and emission wavelengths of approximately 280 nm and 350 nm, respectively [33]. BCA assay is based on the formation of a Cu^{2+}-protein complex in alkaline environments. Peptide bonds, Trp, Tyr and Cys residues are able to reduce the Cu^{2+} to Cu^+, forming a purple-blue complex with BCA which absorbs at 562 nm. BCA assay presents a lower sensitivity when compared to the quantification by tryptophan fluorescence emission, which can be observed in Figure 2.13. Therefore, for simplicity and relevance, only fluorescence results are presented here.

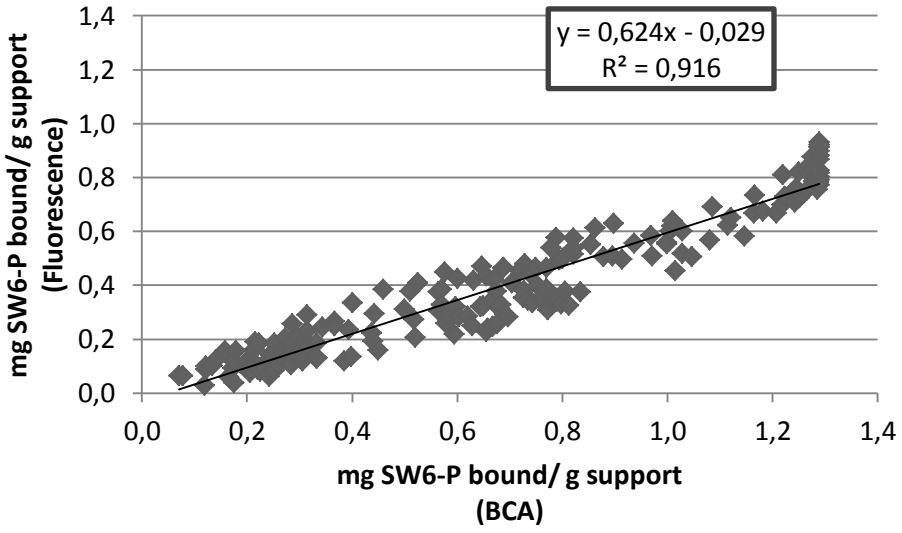


Figure 2.13 – Amount (in mg) of SW6-P peptide bound per g of ligand-functionalized agarose (either with Petasis-Ugi, Ugi or triazine scaffolds). The fact that the slope of the trendline is lower than 1 means that for the same ligand the amount of non-bound peptide detected is higher when using tryptophan fluorescence than BCA assay, which means the fluorescence method is more sensitive. This figure was plotted with data from 188 ligands.

The binding capacity is given by the amount of peptide (either SW6 or SW6-P) bound per gram of ligand-functionalized agarose (Equation 2.1). The results were also analysed in terms of SW6-P mole fraction, in order to determine the selectivity of the ligands towards the phosphorylated peptide (Equation 2.2). The ligands are only selective towards phosphopeptides if the phosphopeptide mole fraction is above 50%.

Equation 2.1 - Ligand binding capacity in mg of peptide per g of support.

$$Binding\ Capacity\ (mg/g) = \frac{mg\ peptide\ bound}{g\ support}$$

Equation 2.2 - Phosphopeptide mole fraction (%).

$$Phosphopeptide\ Mole\ Fraction\ (\%) = \frac{\frac{mol\ phosphopeptide}{g\ support}}{\frac{mol\ phosphopeptide}{g\ support} + \frac{mol\ non-phosphopeptide}{g\ support}} x100$$

Petasis-Ugi ligands present low binding capacities both at pH 4 and pH 8, probably due to the low amount of ligand existent in the beads, as the Petasis reaction presented low yields as previously stated. In both cases, there was a maximum of 30% SW6-P peptide bound. At pH 4 all the ligands that bind to SW6-P are selective towards this peptide, presenting SW6-P mole fractions of 100%. The other ligands do not bind at all to either SW6-P or SW6, which might be happening due to the lack of affinity, although we cannot rule out the possibility of synthesis failure and, therefore, not having the desired ligand or having it at low concentrations. A larger number of ligands bind at pH 8, but the maximum binding capacity and SW6-P mole fraction values observed are similar to the ones obtained at pH 4 (Figure 2.14). Amines A1, A2, A3 and aldehyde B7 seem to be playing key roles on SW6-P binding at pH 4, while amine A4 and aldehyde B7 seem to be critical for binding at pH 8.

Ugi ligands present much higher binding capacities than Petasis-Ugi ligands, both at pH 4 and pH 8. However, by the SW6-P mole fractions it is clear that the selectivity for the phosphopeptide is higher at pH 4. At this pH, all ligands presenting higher binding capacities comprise amine A8, which mimics His residue, suggesting that the imidazole group may be playing a decisive role in phosphate coordination. Generally speaking, the carboxylic acid component does not seem to be critical for binding, since ligands with the same amine and different carboxylic acids have similar binding profiles (Figure 2.15).

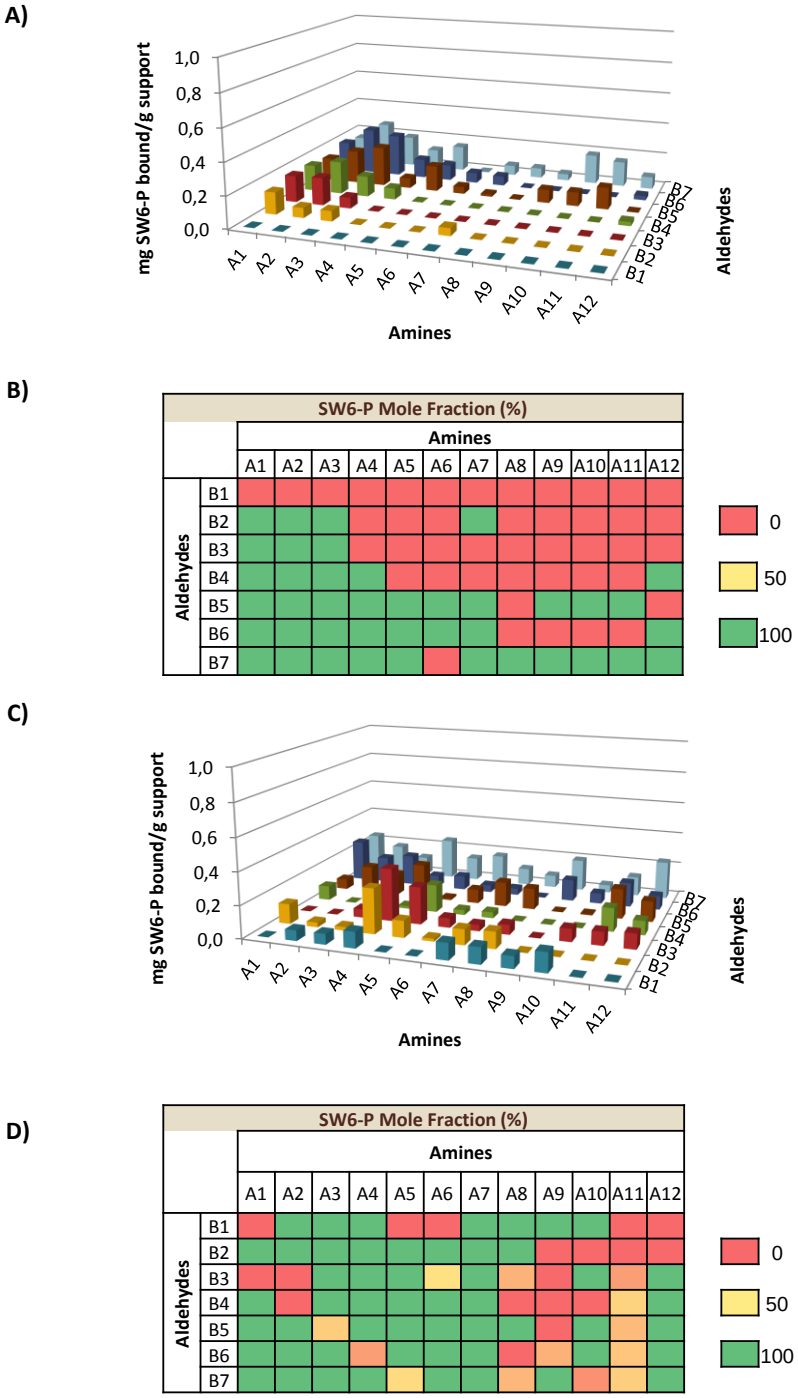


Figure 2.14 – Screening results of Petasis-Ugi ligands. Screenings were performed by adding 0.25 mL of peptide in binding buffer (1 mg/ mL) to each well of a 96-well filter plate containing 0.25 g of ligand-functionalized resin. The unbound peptide from flow-through and washes was collected in 96-well black microplates and the fluorescence intensity of the samples was measured in a microplate reader (λexc = 280 (20) nm – λem = 340 (35) nm filter). A) Amount of SW6-P peptide bound per g of ligand-functionalized support using 50 mM acetate buffer pH 4 as binding buffer. B) Schematic diagram representing the SW6-P mole fraction (%) determined when using 50 mM acetate buffer pH 4 as binding buffer. C) Amount of SW6-P peptide bound per g of ligand-functionalized support using 50 mM ammonium bicarbonate buffer pH 8 as binding buffer. D) Schematic diagram representing the SW6-P mole fraction (%) determined when using 50 mM ammonium bicarbonate buffer pH 8 as binding buffer.

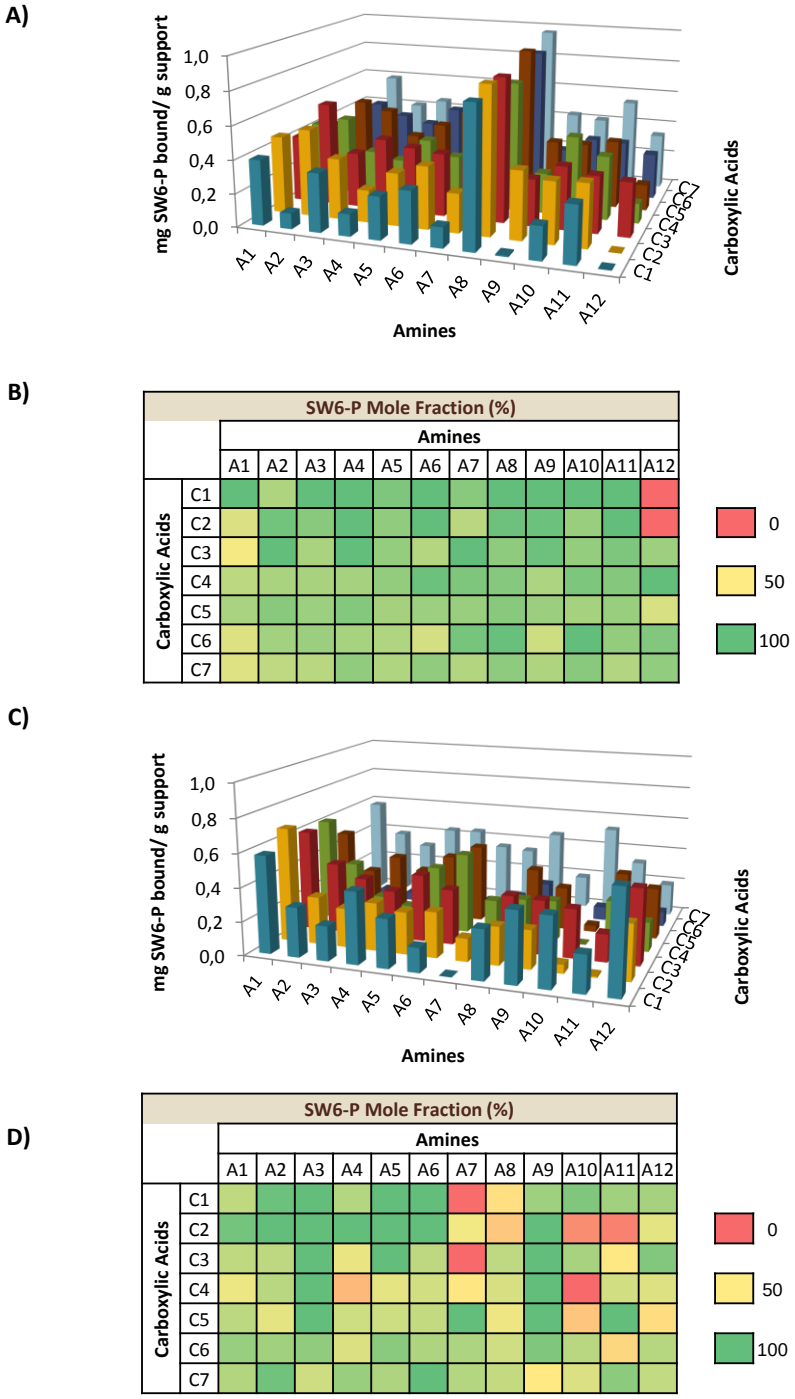


Figure 2.15 – Screening results of Ugi ligands. Screenings were performed by adding 0.25 mL of peptide in binding buffer (1 mg/ mL) to each well of a 96-well filter plate containing 0.25 g of ligand-functionalized resin. The unbound peptide from flow-through and washes was collected in 96-well black microplates and the fluorescence intensity of the samples was measured in a microplate reader (λexc = 280 (20) nm – λem = 340 (35) nm filter). A) Amount of SW6-P peptide bound per g of ligand-functionalized support using 50 mM acetate buffer pH 4 as binding buffer. B) Schematic diagram representing the SW6-P mole fraction (%) determined when using 50 mM acetate buffer pH 4 as binding buffer. C) Amount of SW6-P peptide bound per g of ligand-functionalized support using 50 mM ammonium bicarbonate buffer pH 8 as binding buffer. D) Schematic diagram representing the SW6-P mole fraction (%) determined when using 50 mM ammonium bicarbonate buffer pH 8 as binding buffer.

Triazine ligands present both higher binding capacities and higher SW6-P mole fractions at pH 4. This might be related not only with the amines used for the ligands, but also with the triazine ring itself, since one of its nitrogen atoms may be protonated at acidic pH, and might also be involved in phosphate coordination (Figure 2.16).

All libraries presented higher percentages of binding and higher selectivity towards SW6-P at pH 4; therefore, this pH was chosen to continue the studies. As the initial libraries were only screened once, it is fundamental to re-synthesize and re-screen the best ligands in order to determine the reproducibility of the results. The criteria used to select the best ligands were different for each library and are presented in Table 2.4. The ligands were not only chosen based on these criteria, but also on their structures. Therefore, the 16 Petasis-Ugi ligands selected were narrowed to 7, since some of the ligands did not display relevant structural differences. Conversely, 2 additional Ugi ligands (A1C2 e A1C6) were selected to join the 8 presented on Table 2.4. Therefore, a total of 7 Petasis-Ugi ligands, 10 Ugi ligands, and 10 triazine ligands were re-synthesized and re-screened at pH 4. The screening results resultant from three synthesis and screenings for Petasis-Ugi, Ugi and triazine ligands are presented in Figure 2.17, Figure 2.18, and Figure 2.19, respectively. In general terms, ligands with low binding capacities present high error bars, which is observable for all 7 Petasis-Ugi ligands, and for some Ugi (A1C1, A1C2, A1C7, A2C3, A4C3) and triazine (A1A3, A5A4, A7A3, and A7A4) ligands. This might be happening due to two main factors: (i) low ligand density, which makes the accessibility of the ligand to the target peptide restricted; (ii) the affinity of the ligand to the SW6-P target is low and the binding is mainly governed by non-specific interactions. The rest of triazine and Ugi ligands present error bars below 26% and 18% on average, respectively. Ligand A3B6 is the lead Petasis-Ugi ligand, presenting the higher binding capacity and the lowest error bar (27%).

Noteworthy, amine A8 is present in 4 triazine ligands and 5 Ugi ligands. Moreover, the 5 amino acids (Arg (A7), Lys (A5), Tyr (A4), Ser (A2), His (A8)) which were identified in the structural studies as key players in phosphate recognition are all represented in triazine lead ligands (A1A8, A5A4, A6A2, A6A4, A7A3, A7A4, A8A3, A8A4, A8A6).

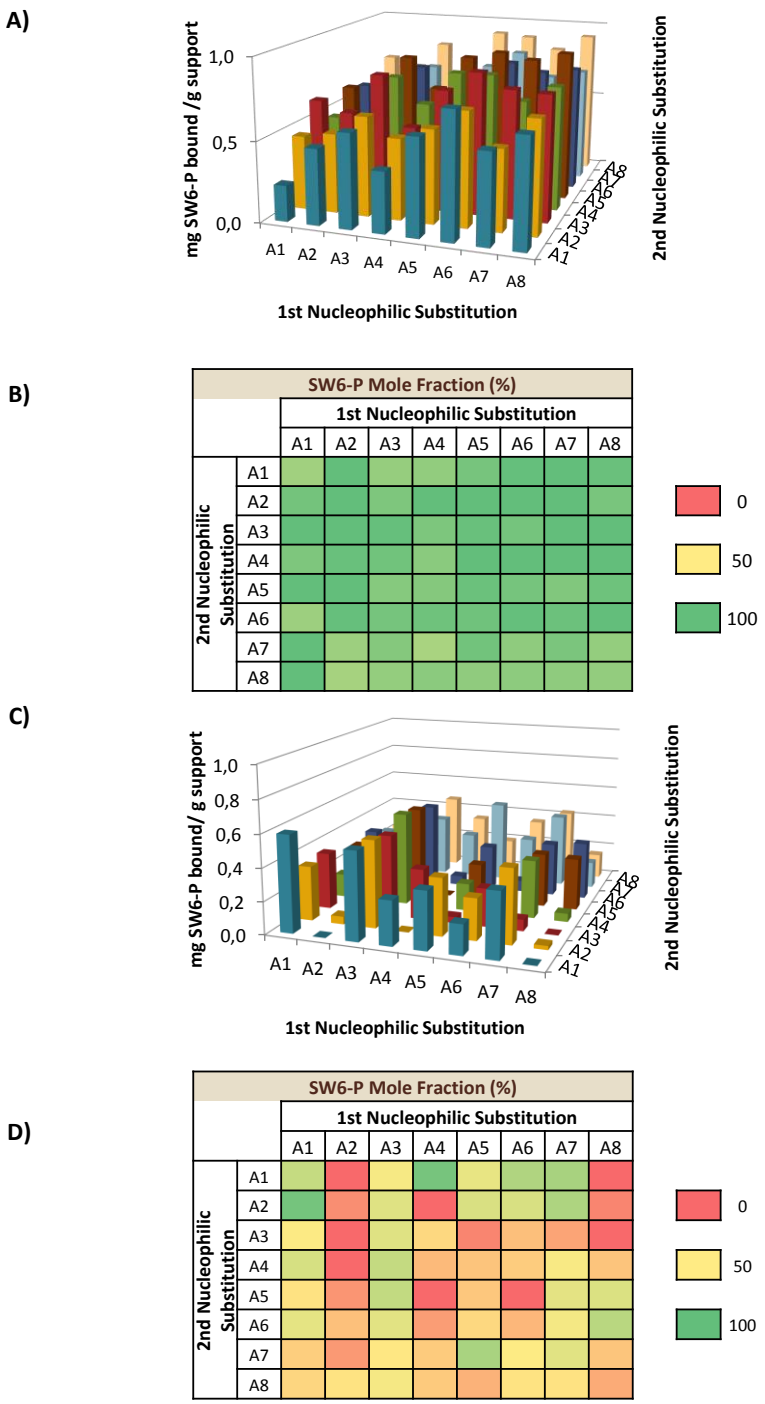


Figure 2.16 – Screening results of triazine ligands. Screenings were performed by adding 0.25 mL of peptide in binding buffer (1 mg/ mL) to each well of a 96-well filter plate containing 0.25 g of ligand-functionalized resin. The unbound peptide from flow-through and washes was collected in 96-well black microplates and the fluorescence intensity of the samples was measured in a microplate reader (λexc = 280 (20) nm – λem = 340 (35) nm filter). A) Amount of SW6-P peptide bound per g of ligand-functionalized support using 50 mM acetate buffer pH 4 as binding buffer. B) Schematic diagram representing the SW6-P mole fraction (%) determined when using 50 mM acetate buffer pH 4 as binding buffer. C) Amount of SW6-P peptide bound per g of ligand-functionalized support using 50 mM ammonium bicarbonate buffer pH 8 as binding buffer. D) Schematic diagram representing the SW6-P mole fraction (%) determined when using 50 mM ammonium bicarbonate buffer pH 8 as binding buffer.

Table 2.4 – Criteria used for the selection of best ligands.

Ligand	% Binding SW6-P	SW6-P Mole Fraction (%)	Number of ligands
Petasis-Ugi	≥ 15	100	16
Ugi	> 30	100	8
	> 60	90 ≥ x > 100	
Triazine	≥ 50	100	10

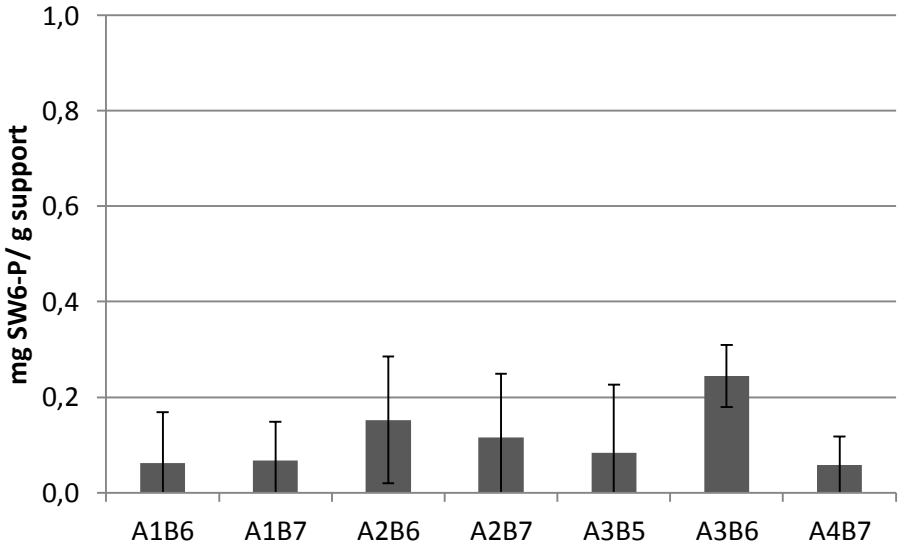


Figure 2.17 – Screening results of Petasis-Ugi ligands (N=3). Screenings were performed by adding 0.25 mL of peptide in 50 mM acetate pH 4 (1 mg/ mL) to each well of a 96-well filter plate containing 0.25 g of ligand-functionalized resin. The unbound peptide from flow-through and washes was collected in 96-well black microplates and the fluorescence intensity of the samples was measured in a microplate reader (λexc = 280 (20) nm – λem = 340 (35) nm filter).

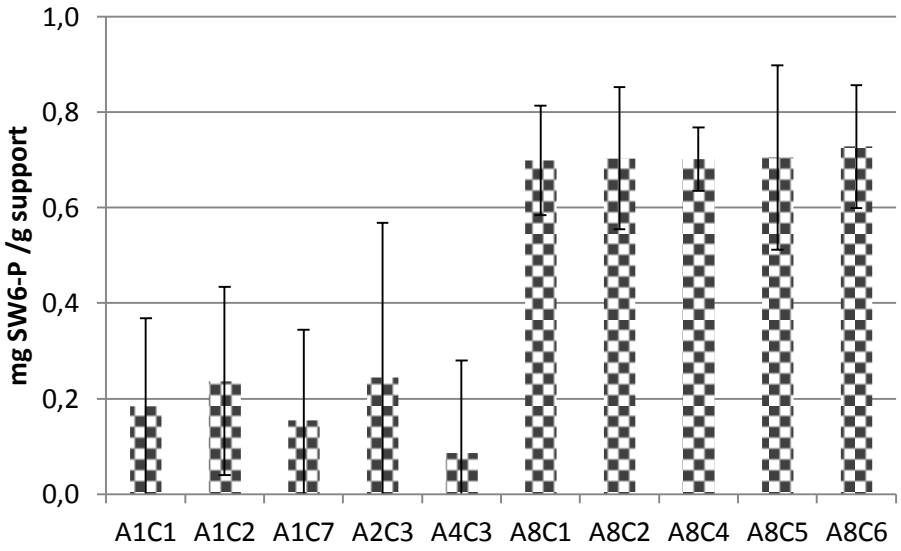


Figure 2.18 – Screening results of Ugi ligands (N=3). Screenings were performed by adding 0.25 mL of peptide in 50 mM acetate pH 4 (1 mg/ mL) to each well of a 96-well filter plate containing 0.25 g of ligand-functionalized resin. The unbound peptide from flow-through and washes was collected in 96-well black microplates and the fluorescence intensity of the samples was measured in a microplate reader (λexc = 280 (20) nm – λem = 340 (35) nm filter).

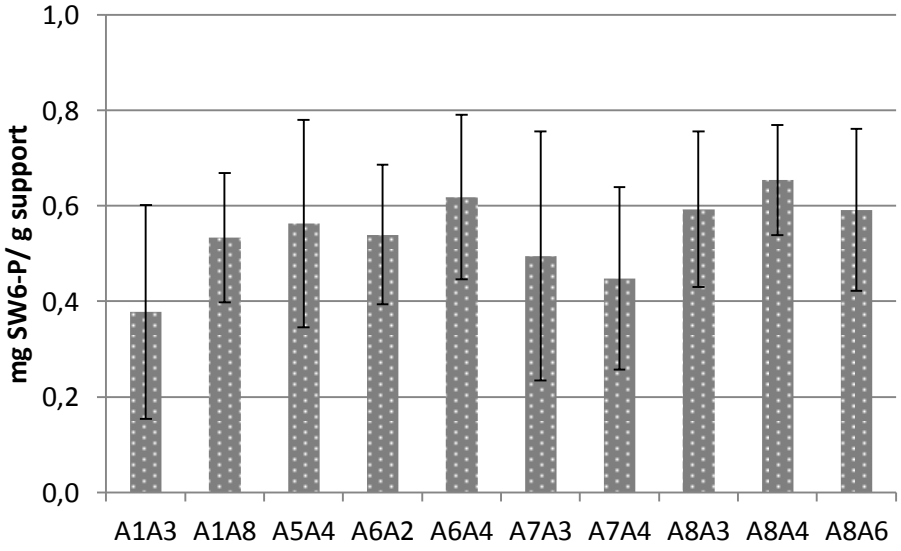


Figure 2.19 – Screening results of triazine ligands (N=3). Screenings were performed by adding 0.25 mL of peptide in 50 mM acetate pH 4 (1 mg/ mL) to each well of a 96-well filter plate containing 0.25 g of ligand-functionalized resin. The unbound peptide from flow-through and washes was collected in 96-well black microplates and the fluorescence intensity of the samples was measured in a microplate reader (λexc = 280 (20) nm – λem = 340 (35) nm filter).

2.4. Conclusions

In this work, by performing structural studies of the natural occurring human PBDs, three different libraries of biomimetic ligands consisting of 232 small molecules were developed. These ligands were synthesized in a solid-phase combinatorial high-throughput platform using three different chemical scaffolds, based on triazine reaction, the Ugi reaction and the Petasis-Ugi reactions. The feasibility of these reactions on agarose was proved by ICP-AES and fluorescence microscopy assays. Ligands were screened against the SW6-P peptide, which comprises the consensus sequence of the BRCT domain, and its non-phosphorylated counterpart at pH 4 and pH 8. The use of acidic conditions revealed to be beneficial, leading to higher binding capacities and phosphopeptide mole fractions. Lead ligands comprised components mimicking His, Tyr, Arg, Ser and Lys amino acid residues, which are systematically observed in the coordination of phosphate to PBDs.

2.5. Bibliography

1. Roque, A.C.A. and C.R. Lowe, *Advances and applications of de novo designed affinity ligands in proteomics.* Biotechnology Advances, 2006. **24**(1): p. 17-26.
2. Roque, A.C., C.S. Silva, and M.Â. Taipa, *Affinity-based methodologies and ligands for antibody purification: Advances and perspectives.* Journal of Chromatography A, 2007. **1160**(1): p. 44-55.
3. Lowe, C.R., S.J. Burton, N.P. Burton, W.K. Alderton, J.M. Pitts, and J.A. Thomas, *Designer dyes:'biomimetic'ligands for the purification of pharmaceutical proteins by affinity chromatography.* Trends in Biotechnology, 1992. **10**: p. 442-448.
4. Roque, A.C.A., M.Â. Taipa, and C.R. Lowe, *Synthesis and screening of a rationally designed combinatorial library of affinity ligands mimicking protein L from Peptostreptococcus magnus.* Journal of Molecular Recognition, 2005. **18**(3): p. 213-224.
5. Teng, S.F., K. Sproule, A. Hussain, and C.R. Lowe, *A strategy for the generation of biomimetic ligands for affinity chromatography. Combinatorial synthesis and biological evaluation of an IgG binding ligand.* Journal of Molecular Recognition, 1999. **12**(1): p. 67-75.
6. Palanisamy, U.D., A. Hussain, S. Iqbal, K. Sproule, and C.R. Lowe, *Design, synthesis and characterisation of affinity ligands for glycoproteins.* Journal of Molecular Recognition, 1999. **12**(1): p. 57-66.
7. Sproule, K., P. Morrill, J.C. Pearson, S.J. Burton, K.R. Hejnæs, H. Valore, S. Ludvigsen, and C.R. Lowe, *New strategy for the design of ligands for the purification of pharmaceutical proteins by affinity chromatography.* Journal of Chromatography B: Biomedical Sciences and Applications, 2000. **740**(1): p. 17-33.
8. Morrill, P.R., G. Gupta, K. Sproule, D. Winzor, J. Christensen, I. Mollerup, and C.R. Lowe, *Rational combinatorial chemistry-based selection, synthesis and evaluation of an affinity adsorbent for recombinant human clotting factor VII.* Journal of Chromatography B, 2002. **774**(1): p. 1-15.

9. Roque, A.C.A., M.Â. Taipa, and C.R. Lowe, *A new method for the screening of solid-phase combinatorial libraries for affinity chromatography.* Journal of Molecular Recognition, 2004. **17**(3): p. 262-267.
10. Roque, A.C.A. and C.R. Lowe (2008) *Affinity Chromatography*, In M. Zachariou (Ed.), Affinity Chromatography, Methods in Molecular Biology, Humana Press Vol. 421, p. 1-23.
11. Liu, R., A.M. Enstrom, and K.S. Lam, *Combinatorial peptide library methods for immunobiology research.* Experimental Hematology, 2003. **31**(1): p. 11-30.
12. Lowe, C.R., *Combinatorial approaches to affinity chromatography.* Current Opinion in Chemical Biology, 2001. **5**(3): p. 248-256.
13. Vetter, S.W. and Z.-Y. Zhang, *Combinatorial chemistry and peptide library methods to characterize protein phosphatases.* Methods in Enzymology, 2003. **366**: p. 260-282.
14. Ruijter, E., R. Scheffelaar, and R.V. Orru, *Multicomponent reaction design in the quest for molecular complexity and diversity.* Angewandte Chemie International Edition, 2011. **50**(28): p. 6234-6246.
15. Lowe, C.R., K. Sproule, R. Li, D.J. Stewart, J.C. Pearson, and S.J. Burton (2000) *Triazine based ligands and use thereof*, U.S. Patent No 6,117,996.
16. Ugi, I., R. Meyr, U. Fetzer, and C. Steinbrückner, *Experiments using isonitrilen.* Angewandte Chemie 1959. **71**: p. 386.
17. Tron, G.C., *Off the Beaten Track: The Use of Secondary Amines in the Ugi Reaction.* European Journal of Organic Chemistry, 2013. **2013**(10): p. 1849-1859.
18. Petasis, N.A. and I. Akritopoulou, *The boronic acid mannich reaction: A new method for the synthesis of geometrically pure allylamines.* Tetrahedron Letters, 1993. **34**(4): p. 583-586.
19. Naskar, D., A. Roy, W.L. Seibel, and D.E. Portlock, *Novel Petasis boronic acid–Mannich reactions with tertiary aromatic amines.* Tetrahedron Letters, 2003. **44**(31): p. 5819-5821.
20. Candeias, N.R., F. Montalbano, P.M.S.D. Cal, and P.M.P. Gois, *Boronic Acids and Esters in the Petasis-Borono Mannich Multicomponent Reaction.* Chemical Reviews, 2010. **110**(10): p. 6169-6193.
21. Petasis, N.A. and I.A. Zavialov, *A New and Practical Synthesis of α-Amino Acids from Alkenyl Boronic Acids.* Journal of the American Chemical Society, 1997. **119**(2): p. 445-446.
22. Candeias, N.R., P.M.S.D. Cal, V. André, M.T. Duarte, L.F. Veiros, and P.M.P. Gois, *Water as the reaction medium for multicomponent reactions based on boronic acids.* Tetrahedron, 2010. **66**(14): p. 2736-2745.
23. Candeias, N.R., R. Paterna, P.M.S.D. Cal, and P.M.P. Góis, *A Sustainable Protocol for the Aqueous Multicomponent Petasis Borono–Mannich Reaction.* Journal of Chemical Education, 2012. **89**(6): p. 799-802.
24. Southwood, T.J., M.C. Curry, and C.A. Hutton, *Factors affecting the efficiency and stereoselectivity of α-amino acid synthesis by the Petasis reaction.* Tetrahedron, 2006. **62**(1): p. 236-242.
25. Kaiser, E., R. Colescott, C. Bossinger, and P. Cook, *Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides.* Analytical Biochemistry, 1970. **34**(2): p. 595-598.
26. Lowe, C.R., A. Hussain, M.L. Mimmack, and J.M. Haigh (2008) *Biomolecule binding ligands*, US Patent Application 12/666,322.
27. Portlock, D.E., D. Naskar, L. West, R. Ostaszewski, and J.J. Chen, *Solid-phase synthesis of five-dimensional libraries via a tandem Petasis–Ugi multi-component condensation reaction.* Tetrahedron Letters, 2003. **44**(27): p. 5121-5124.
28. Trippier, P.C. and C. McGuigan, *Boronic acids in medicinal chemistry: anticancer, antibacterial and antiviral applications.* MedChemComm, 2010. **1**(3): p. 183-198.

29. Regnier, T., F. Berree, O. Lavastre, and B. Carboni, *Solvent-free one-pot four-component synthesis of 2-aminomorpholines. Access to related diaminoalcohols.* Green Chemistry, 2007. **9**(2): p. 125-126.
30. Jourdan, H., G. Gouhier, L. Van Hijfte, P. Angibaud, and S.R. Piettre, *On the use of boronates in the Petasis reaction.* Tetrahedron Letters, 2005. **46**(46): p. 8027-8031.
31. Schlienger, N., M.R. Bryce, and T.K. Hansen, *The Boronic Mannich Reaction in a Solid-Phase Approach.* Tetrahedron, 2000. **56**(51): p. 10023-10030.
32. Yang, C.J., S. Jockusch, M. Vicens, N.J. Turro, and W. Tan, *Light-switching excimer probes for rapid protein monitoring in complex biological fluids.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(48): p. 17278-17283.
33. Lakowicz, J.R. (2009) *Principles of fluorescence spectroscopy* (3rd ed.), Springer, p. 277-278; 308; 529-531.
34. Śmiechowski, M., *Theoretical pKa prediction of O-phosphoserine in aqueous solution.* Chemical Physics Letters, 2010. **501**(1–3): p. 123-129.

Chapter 2 Combinatorial libraries of synthetic ligands for phosphorylated peptides

CHAPTER 3

Chromatographic evaluation and optimization of lead ligands

Summary

Synthetic affinity ligands based on the Petasis-Ugi, Ugi, and triazine reactions have been successfully applied for the enrichment of a phosphoserine (pSer)-containing peptide known to be specifically recognized by the Brca1 C-terminal (BRCT) domain. This chapter explores the potential of previously identified lead ligands to enrich samples containing various phosphorylated peptides. In particular, the enrichment of peptides phosphorylated in threonine (pThr) and tyrosine (pTyr), and also a multi-phosphorylated peptide, was assessed by high-throughput screening (HTS). These phosphorylated peptides are known to specifically bind other phosphoprotein-binding domains (PBDs), such as Src-homology-2 (SH2) and Forkhead-associated (FHA) domains. Binding conditions have been optimized for the best ligands by changing pH and salt concentration. Elution efficiency has been evaluated with different organic solvents, acids, bases and competitor molecules. Partition equilibrium experiments reveal that the binding of all phosphorylated and non-phosphorylated peptides followed the Langmuir adsorption model. Both phosphorylated and non-phosphorylated peptides present low association equilibrium constants (K_A), but phosphorylated peptides have higher maximum binding capacities (Q_{max}), which might be an indication of higher specificity for phosphorylated peptides.

Chapter 3 Chromatographic evaluation and optimization of lead ligands

3.1. Introduction

High-throughput screening (HTS) methodologies in 96-well format have been widely used to screen a variety of compounds, such as drugs, peptides, and affinity ligands [1-3]. Besides its usefulness in the discovery of novel molecules, HTS can also be applied in process development and optimization of chromatographic separations. It is time and cost-effective, as it allows the parallel screening of different experimental parameters, uses small quantities of materials, and eliminates day-to-day variability [1, 4, 5]. Coffman and co-workers reported the use of a 96-well batch-binding miniaturized robotic system, which allows the easy evaluation and tailoring of several factors, such as incubation time, binding and elution conditions, amount of protein and resin, hold-up volume, etc. Batch-binding methodologies yield information on thermodynamics, selectivity, and binding kinetics. In addition, HTS systems can be coupled to automated analytical systems, such as high-gradient liquid chromatography and capillary electrophoresis [1]. The major challenge of HTS methodologies is mainly related to translation of data to scale-up operations, but it provides qualitative results which can be used to guide the development of laboratory- and manufacturing-scale chromatographic processes [1, 4]. Also, this is not particularly problematic when referring to phosphoprotein and phosphopeptide enrichment, as it usually takes place in small scale.

The optimization of binding conditions depends on the mode of chromatography (e.g. hydrophobic, ion-exchange, mixed-mode, immobilized metal affinity chromatography (IMAC)), involving the proper selection of the buffering species, and the adjustment of experimental parameters, such as pH, temperature, ionic strength, etc. The influence of salt is particularly complex. Hofmeister discovered that protein solubility is not only influenced by the concentration of salt, but also by its nature. He ranked different ions according to their salting-in and salting-out effectiveness for globular proteins (Figure 3.1) [6, 7]. The effects of anions are more pronounced when compared to cations, due to their asymmetric effect on polarizable water molecules [8]. These ions have been classified according to their ability to change the hydrogen bonding network of water. Ions coined kosmotropes are considered to be water-structure makers, having stabilizing and salting-out effects on proteins. Chaotropes are water-structure breakers, and have destabilizing and salting-in effects [6, 9]. However, there is some debate in the scientific community involving this subject. In fact, the solute itself may have influence on the salting-in/ salting-out behaviour, rather than the ion-water interaction alone [9, 10].

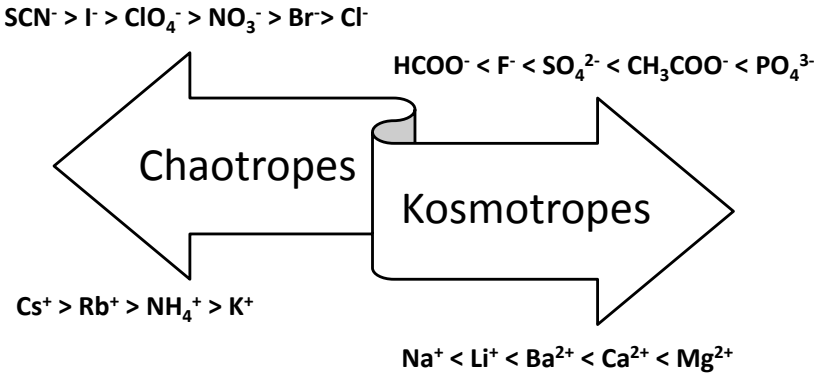


Figure 3.1 – The Hofmeister series. Adapted from [8].

An efficient elution strategy is not straightforward, as simple protocols usually involve incomplete breaking of the protein-ligand interaction. Low pH buffers are commonly used to break ionic bonds, but high pH buffers are also recurrent, and sometimes both are used subsequently. The addition of high concentrations of salt also disrupts ionic bonds, but might foment hydrophobic interactions. Hydrogen bonds are affected by the addition of denaturating agents and detergents, such as urea and guanidine-HCl, but their applicability is limited by the loss of function of the purified protein. Other strategies involve the use of water-miscible organic solvents and competitive elution. The last one has several advantages, namely the use of mild conditions and high specificity, but might have limitations associated with the cost of the competitor molecule and time-effectiveness [11].

3.2. Materials and Methods

3.2.1. Chemicals

All reagents were of the highest purity available and the solvents were pro-analysis.

4-imidazolecarboxylic acid (C1), amino-2-propanol (A2), ammonium bicarbonate (NH_4HCO_3), ammonium hydroxide solution (NH_4OH), boric acid (H_3BO_3), cyanuric chloride, epichlorohydrin, glyoxylic acid monohydrate, histamine (A8), isopropyl isocyanide, N,N-Dimethyl-4,4'-azodianiline (A1), phenethylamine (A3), phenylboronic acid, phosphoric acid (H_3PO_4), propionaldehyde, sodium bicarbonate ($NaHCO_3$), sodium periodate, sodium thiosulfate,

succinamic acid (C2), tetrahydrofuran-3-carboxaldehyde solution (B6), tris(hydroxymethyl)aminomethane, trifluoroacetic acid (TFA), tryptamine (A6), and tyramine (A4) were acquired from Sigma-Aldrich. 2-Propanol and methanol (MeOH) were obtained from Roth, glycine from Acros Organics, and acetic acid glacial from Pronalab. Acetone and dimethylformamide (DMF) were purchased from VWR. Acetic anhydride, acetonitrile (ACN), ethanol absolute PA, hydrochloric acid (HCl) 37%, sodium acetate, sodium chloride (NaCl) and sodium hydroxide (NaOH) were acquired from Panreac. Ser-Gln-Val-Phe-Pro-Trp (SW6), pSer-Gln-Val-Phe-Pro-Trp (SW6-P), Thr-Gln-Val-Asp-Ala-Trp (TW6), pThr-Gln-Val-Asp-Ala-Trp (TW6-P), Tyr-Glu-Glu-Ile-Pro-Trp (YW6), pTyr-Glu-Glu-Ile-Pro-Trp (YW6-P), Tyr-Ala-Gly-Ser-Thr-Asp-Glu-Asn-Thr-Asp-Ser-Glu-Trp (YW13), and Tyr-Ala-Gly-pSer-pThr-Asp-Glu-Asn-pThr-Asp-Ser-Glu-Trp (YW13-P) peptides were > 98% pure and were purchased from Genecust and Caslo.

3.2.2. Chromatographic materials

Sepharose™ CL-6B was acquired from GE Healthcare. Captiva 96-well 20u filter plate with the respective duo seal 96-well Pl seal and Captiva 96-well plate cover were purchased from Agilent Technologies. Half-area UV-Star® 96-well microplates, black immunograde 96-well microplates, and 96-well transparent microplates were obtained from Greiner Bio-One, Brand, and Sarstedt, respectively.

3.2.3. Instrumentation

Ligand synthesis was carried out in an IKA KS 4000 ic control shaker. Fluorescence readings in 96-well microplates were conducted in a Tecan F200 Microplate Reader using λexc = 280 (20) nm – λem = 340 (35) nm; 560 (10) nm; and 280 (5) nm filters (Tecan).

3.2.4. Software

Ligands and peptides were sketched on ChemBioDraw® Ultra 13.0 software, and their properties were determined using charge, logD, major microspecies, and p*K*a plugins from MarvinSketch 4.1.13 software (ChemAxon). Data from partition equilibrium experiments was fitted by nonlinear regression using OriginPro 8.5.1 SR2.

3.2.5. Methods

3.2.5.1. Solid-phase synthesis of Petasis-Ugi ligands

Petasis ligand was synthesized on agarose according to §2.2.4.5 and §2.2.4.6. Petasis-modified agarose was washed with MeOH (20% to 100% in 20% increments) in a sinter funnel, and the resin was weighted and equally distributed between 2 flasks. In the meantime, each amine and aldehyde components (5 molar eq. of each relative to epoxy) were dissolved in MeOH (0.5 mL/ g moist resin / component/ flask) and were reacted for 2h at 60°C with agitation (200 rpm). Each mixture of amine and aldehyde was then added to the Petasis-modified agarose (1 mL/ g moist resin/ flask), followed by isopropyl isocyanide (5 molar eq. relative to epoxy). The reaction occurred for 48h at 60°C with constant shaking (200 rpm). At the end, Petasis-Ugi ligands were washed with MeOH (100% to 20% in 20% decrements) and finally with distilled deionised (dd) water (10x resin volume). The amine components were A2 and A3, while the aldehyde component was B6.

3.2.5.2. Solid-phase synthesis of Ugi ligands

Aldehyde-functionalized agarose was prepared according to §2.2.4.8, washed with MeOH in 20% increments using vacuum filtration, weighted and distributed among three flasks. Amine components (5 molar eq. of each relative to epoxy) were dissolved either in MeOH (for A1) or 50% (v/v) DMF/ MeOH (for A8) (0.5 mL/ g moist resin/ component/ flask), and added to each flask containing aldehyde-modified agarose. The mixture was incubated for 2h at 60°C in an orbital shaker (200 rpm). Each carboxylic acid component (5 molar eq. of each relative to epoxy) was then added to the mixture (0.5 mL/ g moist resin/ component/ flask), followed by isopropyl isocyanide (5 molar eq. relative to epoxy), and the reaction was carried on at 60°C for 48h with agitation (200 rpm). Lastly, Ugi ligands were washed with MeOH in 20% decrements, followed by dd water (10x resin volume). The carboxylic acids used were C1 and C2; the first one was dissolved in 50% (v/v) DMF/ MeOH, and the second in MeOH.

3.2.5.3. Solid-phase synthesis of triazine ligands

Cyanuric chloride-modified agarose was prepared according to §2.2.4.10, thoroughly washed with dd water, weighted and divided between 4 flasks. The first amine component (2 molar eq. relative to epoxy; 1 mL/ g moist resin/ flask) was added to each flask and the reaction occurred for 24h at 30°C. After this period, the ligands were washed with the solvent in which each amine was dissolved (5x resin volume) and dd water (5x resin volume). The second amine component (5 molar eq. relative to epoxy; 1 mL/ g moist resin/ flask) was then added to each flask, and the mixture was reacted for 48h at 80°C. In the end, the ligands were washed with the solvent in which each amine was solubilized (5x resin volume), 0.1 M HCl (1x resin volume), dd water (1x resin volume), 0.1 M NaOH in 30% isopropanol (v/v) (regeneration buffer; 1x resin volume), and finally dd water (5x resin volume). The amine components were A1, A3, A4, A6, and A8. All the amines were dissolved in dd water, with the exception of A1 and A6, which were dissolved in 50% (v/v) DMF/ dd water. $NaHCO_3$ 1M (1 molar eq. relative to epoxy) was added to each amine.

3.2.5.4. Screening of ligands against phosphorylated and non-phosphorylated peptides

Petasis-Ugi, Ugi, and triazine ligands were distributed between the wells of 96-well filter plates (0.25 g of ligand-functionalized agarose/ well). The plates were washed with regeneration buffer followed by dd water (3 cycles of washes, 0.75 mL/ well), followed by 50 mM acetate pH 4 (8x 0.75 mL/ well on average), until the UV absorbance of the washes at 280 nm was below 0.005 (last washes were collected in a half-area UV-Star® 96-Well microplate). TW6, TW6-P, YW6, and YW6-P peptides were solubilized in 50 mM acetate pH 4 (1 mg/ mL) and 0.25 mL of each peptide was loaded per well. The ligand-functionalized resins were incubated with the peptides for 1h at room temperature with manual shaking. The libraries were then washed with 50 mM acetate pH 4 (8x 0.25 mL). Both the flow-through and washes were collected in 96-well black microplates by centrifugation at 600 rpm for 20 s, and the fluorescence intensity was read on a microplate reader using a λexc = 280 (20) nm – λem = 340 (35) nm filter.

3.2.5.5. Optimization of binding conditions

Binding conditions were optimized for ligands A6A4, A8A3, A8C2, and A3B6. These ligands were distributed by the wells of 96-well filter plates (0.25 g/ well), and incubated with one of the following: (i) a mixture of SW6-P, TW6-P and YW6-P (0.5 mg/mL total peptide; 0.25 mL/ well); (ii) a mixture of SW6, TW6 and YW6 (0.5 mg/mL total peptide; 0.25 mL/ well); (iii) YW13-P (0.5 mg/mL; 0.25 mL/ well); or (iv) YW13 (0.5 mg/mL; 0.25 mL/ well). The binding buffers tested were the following: 50 mM Gly pH 3; 50 mM acetate pH 4; 50 mM acetate pH 5; 50 mM acetate pH 6; 50 mM Tris pH 7; 50 mM NH_4HCO_3 pH 8; 50 mM Tris pH 9; 50 mM Gly pH 10; and 50 mM Gly pH 11. Different concentrations of NaCl (25, 50, 100, 250, 500, 1000 mM) were further tested for 50 mM Gly pH 3 buffer. The screening procedure was similar to §3.2.5.4.

3.2.5.6. Screening of negative controls against phosphorylated peptides

Agarose, aminated agarose (agarose-NH_2) and aldehyde-functionalized agarose (agarose-CHO) were screened against all phosphorylated peptides using 50 mM glycine pH 3. The screening procedure was similar to §3.2.5.4.

3.2.5.7. Optimization of elution conditions

Elution conditions were optimized for ligands A8A3 and A8C2. After screening each peptide or peptide mixture using the best binding condition determined in §3.2.5.5, 0.25 mL of each eluent was loaded into each well of the 96-well filter plate. The elution fraction was collected in 96-well black microplates by centrifugation at 600 rpm for 20 s, and the fluorescence intensity was read on a microplate reade using a λexc = 280 (20) nm – λem = 340 (35) nm filter. This step was repeated until no peptide was detected in the elution fraction. The different eluents tested were: 50% (v/v) ACN/H_2O + 1% acetic acid (Sol 1); 50% (v/v) MeOH/H_2O + 1% acetic acid (Sol 2); MeOH (Sol 3); 50% (v/v) ACN/H_2O + 0.1% NH_4OH (Sol 4); 50% (v/v) ACN/H_2O + 1% H_3PO_4 (Sol 5); 50% (v/v) ACN/H_2O + 0.1% TFA + 1% H_3PO_4 (Sol 6); 50% (v/v) ACN/H_2O + 1% H_3BO_3 (Sol 7); 50% (v/v) ACN/H_2O + 0.1% TFA + 1% H_3BO_3 (Sol 8).

3.2.5.8. Partition equilibrium experiments

A8A3- and A8C2-functionalized agarose was distributed by the wells of a 96-well filter plate (0.25 g/ well). The ligands were then incubated overnight at 30°C and 100 rpm with different concentrations (0.05, 0.1, 0.5, 1, 2.5, 5, 7.5, 10 mg/ mL; 0.25 mL/ well) of one the following: (i) a mixture of SW6-P, TW6-P and YW6-P; (ii) a mixture of SW6, TW6 and YW6; (iii) YW13-P; or (iv) YW13. The flow-through was collected in 96-well black microplates by centrifugation at 600 rpm for 20 s, and the fluorescence intensity was read on a microplate reader using a λexc = 280 (20) nm – λem = 340 (35) nm filter.

3.3. Results and Discussion

3.3.1. Screening of lead ligands against pSer-, pThr-, and pTyr-containing peptides

Lead ligands identified in chapter 2 comprised 7 Petasis-Ugi ligands, 10 Ugi ligands, and 10 triazine ligands. On the basis of their binding capacities, synthesis and screening reproducibility, and structural diversity, 2 Petasis-Ugi ligands (A2B6, A3B6), 3 Ugi ligands (A1C2, A8C1, A8C2), and 4 triazine ligands (A1A3, A1A8, A6A4, A8A3) were selected and screened against pThr- and pTyr-containing peptides. The peptides selected were pThr-Gln-Val-Asp-Ala-Trp (TW6-P) and pTyr-Glu-Glu-Ile-Pro-Trp (YW6-P), which comprise consensus sequences recognized by FHA and Src-like SH2 domains, respectively [12, 13]. The binding assays were performed at pH 4, and the results were compared with the screenings performed against pSer-Gln-Val-Phe-Pro-Trp (SW6-P) (Figure 3.2). The binding capacity values (mg phosphopeptide/ g support) indicate that all ligands preferentially bind to SW6-P, followed by YW6-P, and finally TW6-P. The ligands were also screened against the non-phosphorylated versions of these peptides (SW6, TW6, and YW6) in order to determine the phosphopeptide mole fractions, which presented the following order for all ligands: SW6-P mole fraction > TW6-P mole fraction > YW6-P mole fraction. As the first libraries were screened against the SW6-P peptide, and lead ligands were selected based on the binding profiles of these libraries, it is foreseeable that this peptide might present the best performance in terms of binding capacities and phosphopeptide mole fractions. The fact that the YW6-P peptide presents mole fractions of approximately 50% indicates that the ligands are binding indiscriminately to YW6 and YW6-P and, therefore, the recognition is not occurring through the phosphate moiety.

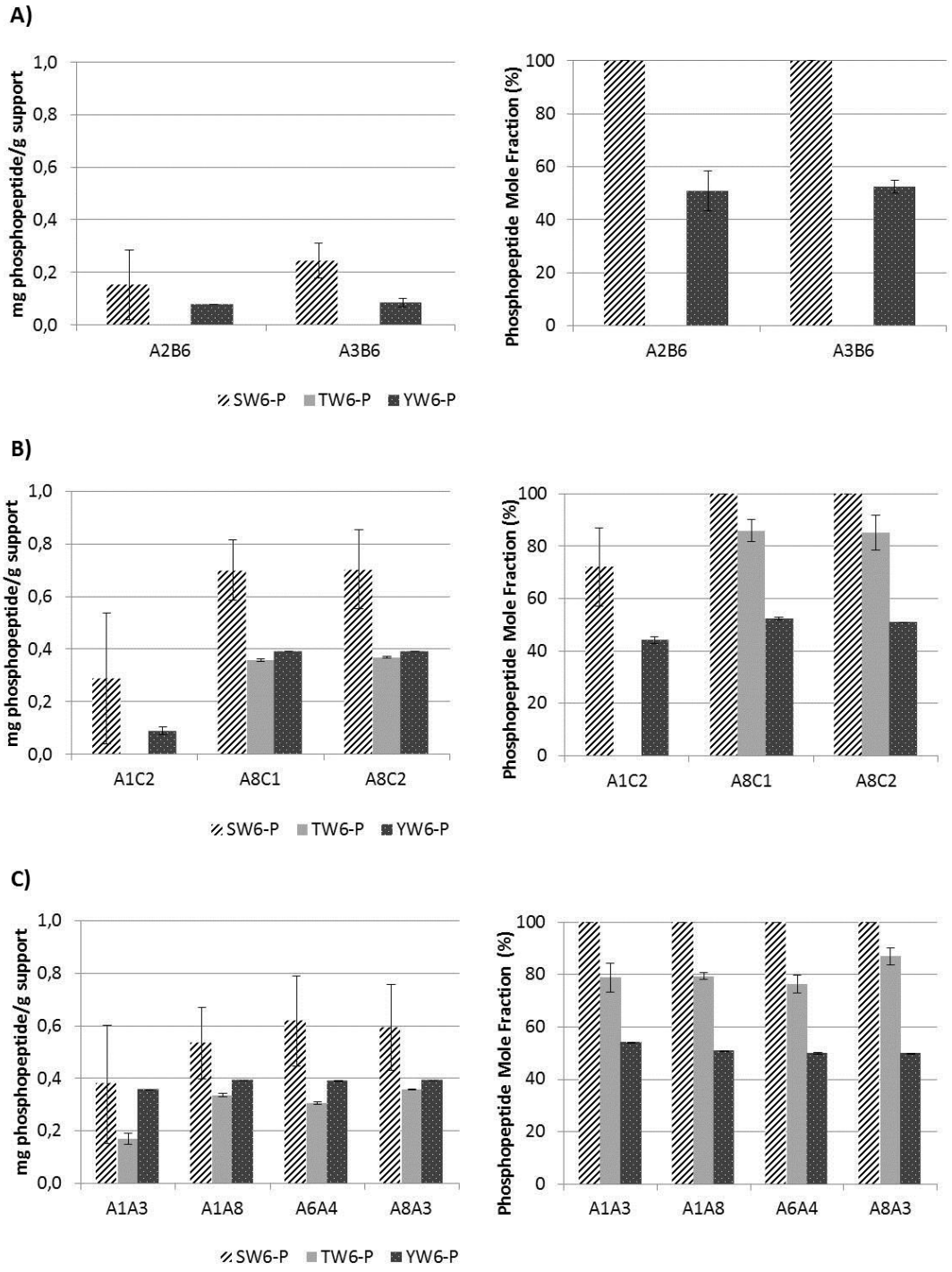


Figure 3.2 – Screening results of lead ligands against SW6-P, TW6-P and YW6-P. Screenings were performed by adding 0.25 mL of peptide in 50 mM acetate buffer pH 4 (1 mg/ mL) to each well of a 96-well filter plate containing 0.25 g of ligand-functionalized resin. The unbound peptide from flow-through and washes was collected in 96-well black microplates and the fluorescence intensity of the samples was measured in a microplate reader (λexc = 280 (20) nm – λem = 340 (35) nm filter). A) Left: Amount of phosphopeptide bound per g of Petasis-Ugi ligand-functionalized support; Right: Phosphopeptide mole fraction (%) for Petasis-Ugi ligands; B) Left: Amount of phosphopeptide bound per g of Ugi ligand-functionalized support; Right: Phosphopeptide mole fraction (%) for Ugi ligands; C) Left: Amount of phosphopeptide bound per g of triazine ligand-functionalized support; Right: Phosphopeptide mole fraction (%) for triazine ligands. (N=2)

To better understand ligand-peptide interactions, all ligands and peptides were first sketched in ChemBioDraw® Ultra 13.0 software, and then their charges and logD at pH 4 were determined using MarvinSketch 4.1.13 software from ChemAxon. LogD is defined as the octanol-water distribution coefficient and may represent compounds at any pH value [14]. This will give an idea of the molecular hydrophobicity of each ligand and peptide at the pH of binding. The structures of Petasis-Ugi based ligands, Ugi-based ligands, and triazine-based ligands are presented in Table 3.1, Table 3.2, and Table 3.3., respectively.

Table 3.1 – Structures of Petasis-Ugi based ligands immobilized on agarose. The white sphere is a schematic representation of an agarose bead.

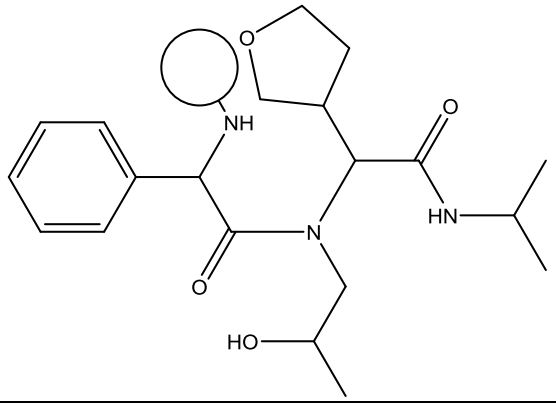
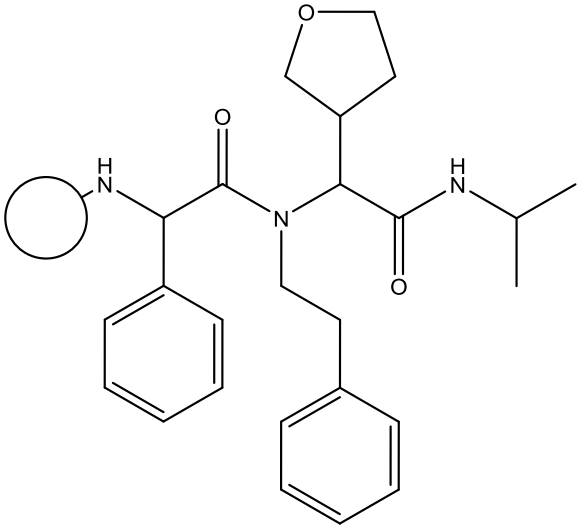
Nomenclature	Structure
A2B6	
A3B6	

Table 3.2 – Structures of Ugi-based ligands immobilized on agarose. The white sphere is a schematic representation of an agarose bead.

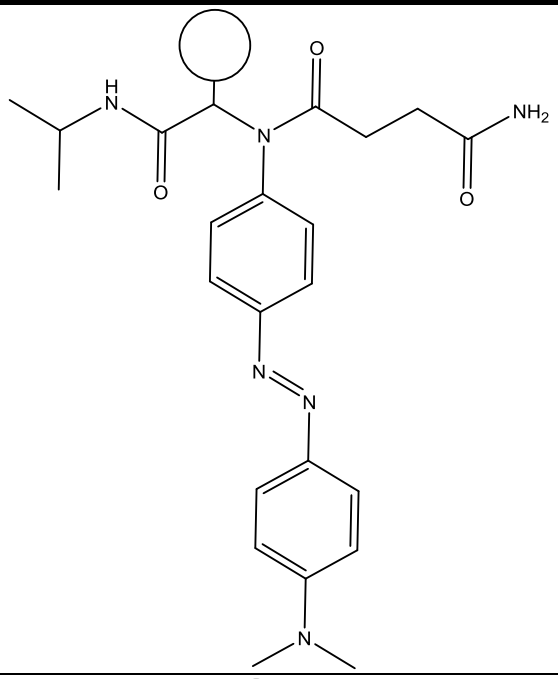
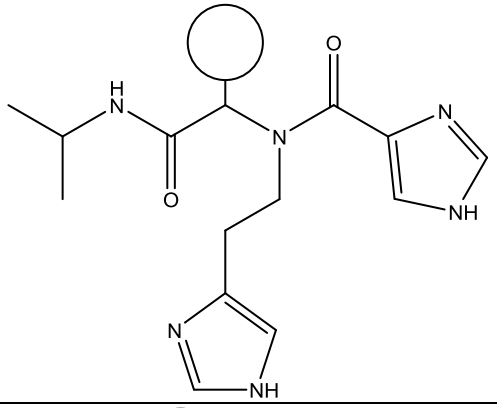
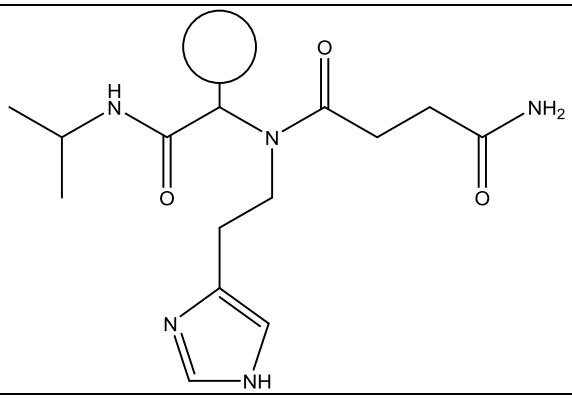
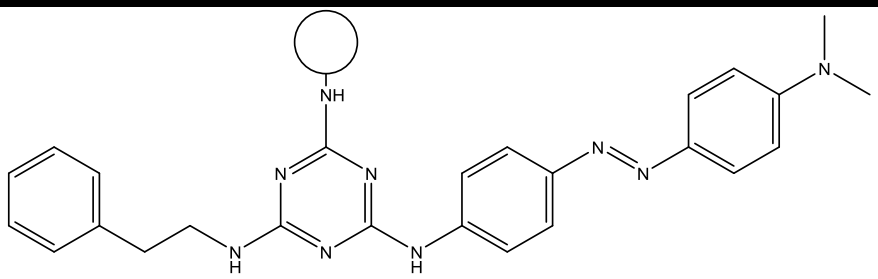
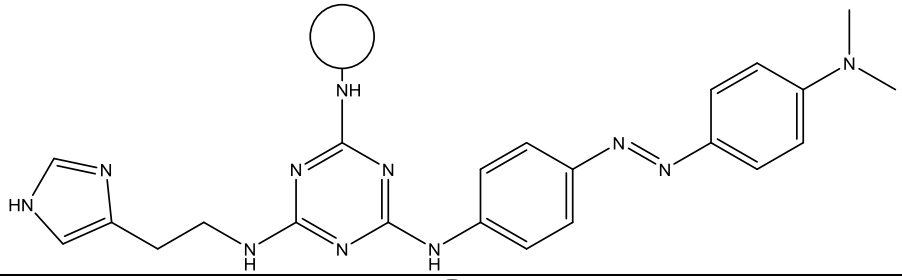
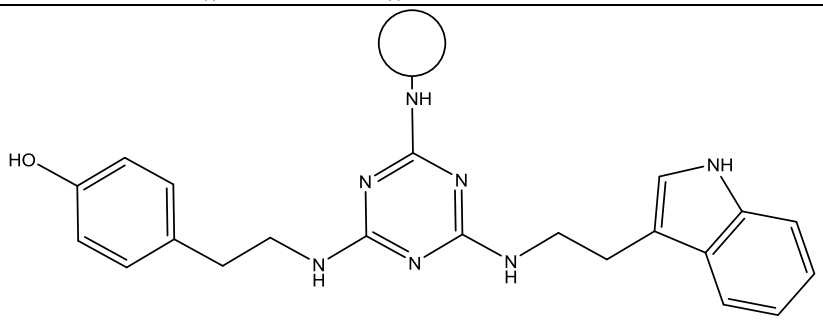
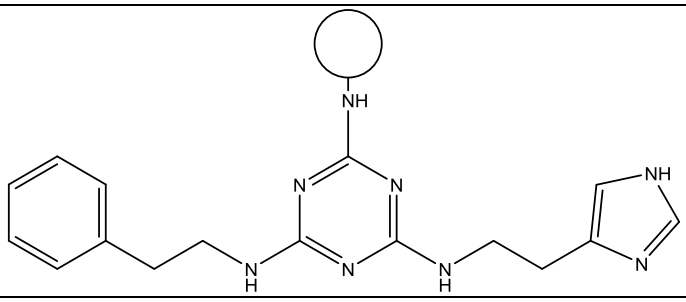
Nomenclature	Structure
A1C2	
A8C1	
A8C2	

Table 3.3 – Structures of triazine-based ligands immobilized on agarose. The white sphere is a schematic representation of an agarose bead.

Nomenclature	Structure
A1A3	
A1A8	
A6A4	
A8A3	

A variety of different interactions may be involved in the ligand-peptide binding, such as hydrogen bonding, hydrophobic interactions, electrostatic interactions, salt-bridging, and van der Waals forces. All lead ligands are positively charged at pH 4, while the peptides present net negative charges, with the exception of SW6 (Figure 3.3).

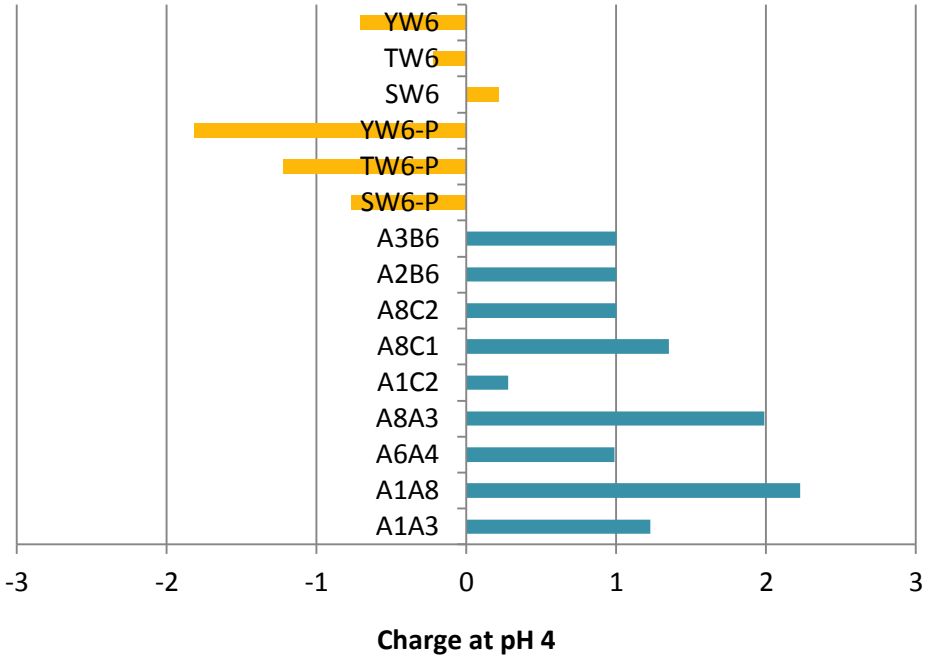


Figure 3.3 – Net charge of ligands and peptides at pH 4. Values were calculated using the Charge Plugin of MarvinSketch software (ChemAxon).

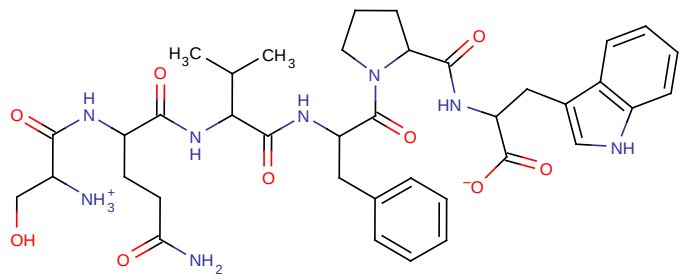
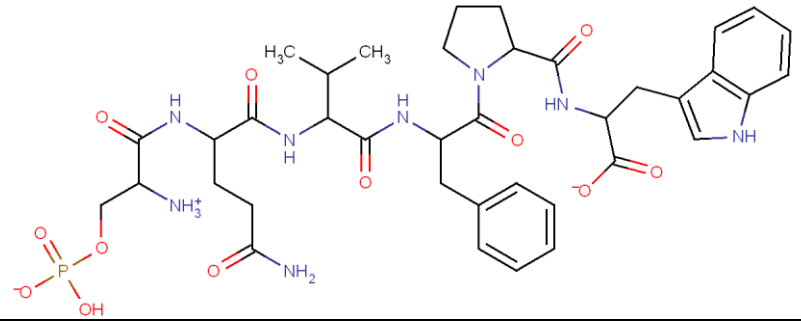
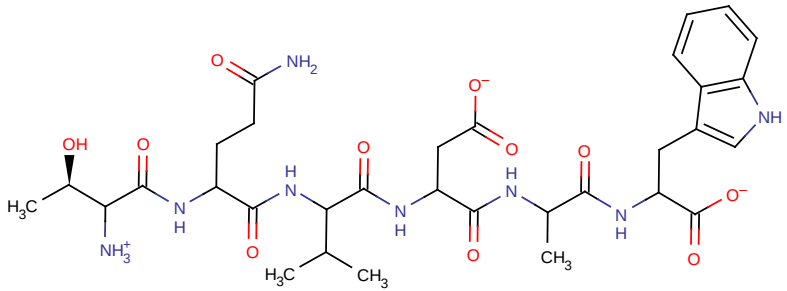
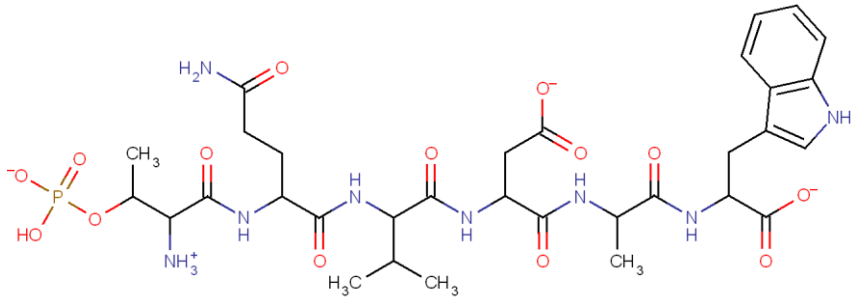
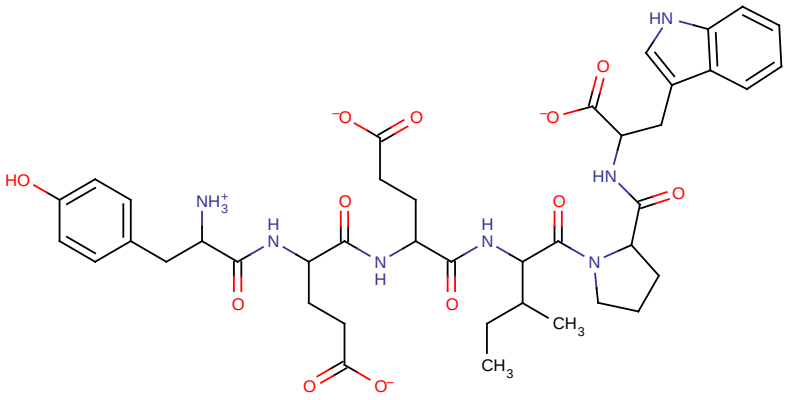
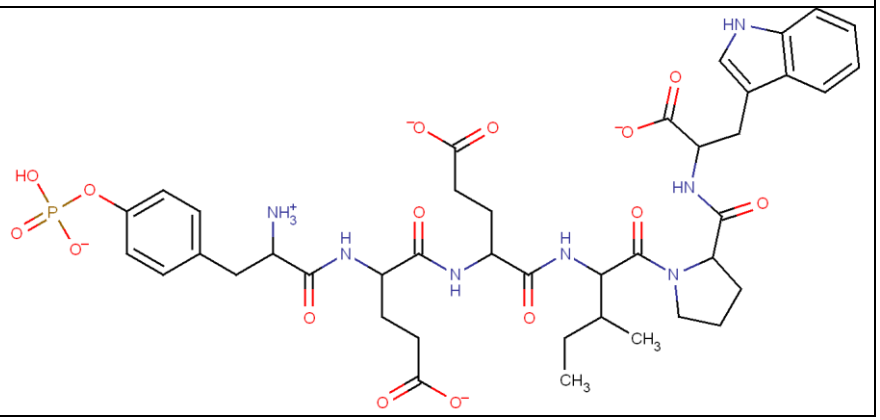
Table 3.4 presents the major protonation form of each peptide at pH 4. Both SW6 and SW6-P have a negative charge on the terminal carboxyl group of the peptide. However, the interaction is apparently occurring mainly through the phosphate group, since SW6 binds 0% to all ligands except A1C2, while SW6-P presents high binding capacities. This is translated into high enrichment efficiencies, which are represented by SW6-P phosphopeptide mole fraction values of 100% (Figure 3.2). Nonetheless, it is plausible that carboxyl groups might interfere in the binding profile of TW6-P and YW6-P peptides, as an increase in the carboxyl group content corresponds to a decrease in the enrichment efficiency. Table 3.4 shows that at pH 4 SW6-P has one negative charge on the terminal carboxyl group, followed by TW6-P with two negative charges on carboxyl groups (one on carboxyl terminal and one on the side chain of Asp residue) and YW6-P with three negative charges (one on carboxyl terminal and two on the side chains of Glu residues). The phosphopeptide mole fractions decrease in the same order, with SW6-P mole fraction > TW6-P mole fraction > YW6-P mole fraction (Figure 3.2). It is reasonable

to consider that the ligands might be binding both to the phosphate and carboxyl groups of TW6-P and YW6-P peptides. Different strategies have been proposed in order to avoid the interference of acidic peptides in phosphopeptide enrichment, with one of the most commonly used protocols being the O-methyl esterification of carboxylic acids [15, 16]. The main limitations of this method are the occurrence of ester hydrolysis; lack of efficiency of the esterification reaction; and the occurrence of side reactions, such as deamidation and subsequent methylation of Asn and Gln amino acid residues [17-19]. Ye and co-workers reported that the use of ACN in loading and washing solutions affects differently the degree of ionization of phosphate and carboxyl groups. The pH values of Asp and Glu amino acids increase for higher ACN concentrations, while the pH of the phosphate moiety remains fairly constant [20].

However, electrostatic interactions are most likely not playing a sole role in the binding efficiency. In fact, although TW6-P has one extra charge in a carboxyl group and a net charge more negative that SW6-P, all ligands present higher binding capacities for SW6-P. The same goes for YW6-P, which is the one with the most negative net charge and which present three negative charges on carboxyl groups. If the only factor to take into consideration would be the net charge of the peptides, then the binding capacities would decrease in the following order: YW6-P > TW6-P > SW6-P, which is not observed. Therefore, it is unlikely that the binding is occurring (or at least exclusively occurring) through the carboxyl groups. Other types of interactions, such as hydrophobic interactions, may be contributing for binding as well. All peptides are hydrophilic at pH 4, which is translated into negative values of logD. However, they all possess amino acid residues able to establish hydrophobic interactions with the ligands, such as Val, Ile, Phe, Trp, etc. YW6 and YW6-P peptides present similar values of logD at pH 4, which means they have similar hydrophobicity (Figure 3.4). On the other hand, YW6-P is more negatively charged than YW6, due to the presence of the phosphate group. In spite of this, ligands bind equally to both peptides. One possibility is that the aromatic group of Tyr may be interacting with the ligands either or both by hydrophobic interactions and amino-aromatic bonding. The aromatic ring of Tyr may be acting as a hydrogen bond acceptor, interacting with a hydrogen bond donor (such as hydroxyl or amine groups) of the ligands. Small partial charges on carbon and hydrogen atoms of the ring are the leading contributors for the interaction, rather than delocalized electrons. The amino-aromatic hydrogen bond presents half the strength of a common hydrogen bond [21, 22]. This type of interaction has been observed in the recognition of pTyr by SH2 domains [23].

A comprehensive understanding of these ligand-peptide interactions would implicate conducting a more detailed structural analysis by nuclear magnetic resonance (NMR) and/ or X-ray crystallography. However, this would require the solution-phase synthesis of the ligands. Until now, the use of solid-state NMR for the structural analysis of organic compounds and biomolecules has been somehow restricted by its limited sensitivity and resolution. However, recent advances in solid-state NMR made possible to detect proteins in the nanomolar range, opening great prospects for ligand characterization and the study of molecular interactions on solid-phase [24, 25].

Table 3.4 – Major protonation forms of SW6, SW6-P, TW6, TW6-P, YW6, and YW6-P peptides at pH 4. The protonation forms were determined using the Major Microspecies Plugin from MarvinSketch (ChemAxon).

Amino acid	Non-phosphorylated	Phosphorylated
Ser		
Thr		
Tyr		

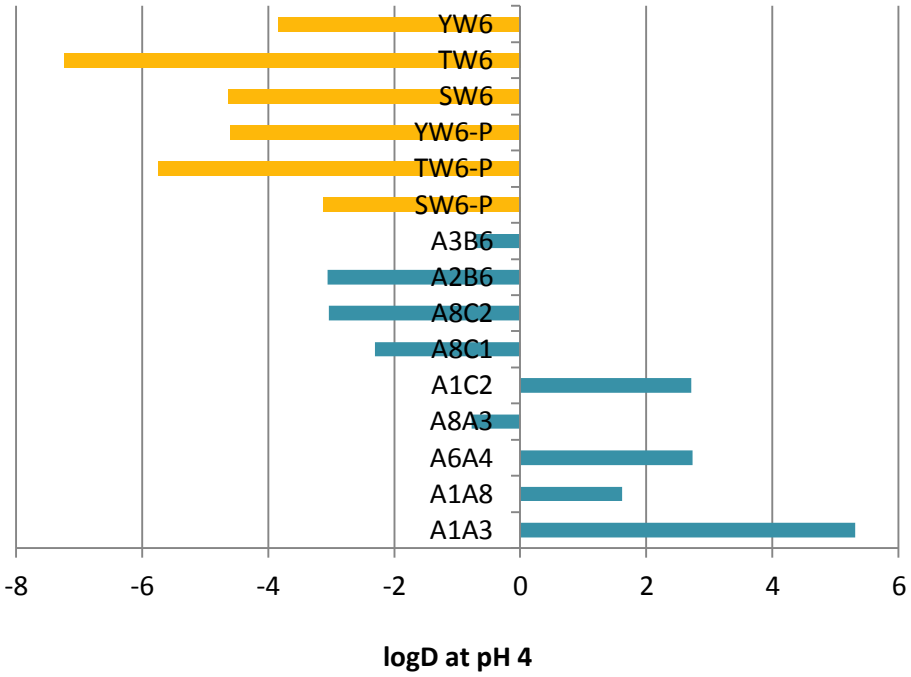


Figure 3.4 – Octanol-water distribution coefficient (logD) of ligands and peptides at pH 4. The values were obtained using logD Plugin of MarvinSketch software (ChemAxon). Hydrophilic compounds present negative logD values, while hydrophobic compounds present positive logD values.

3.3.2. Optimization of binding conditions

Binding conditions were optimized for the best ligands of each library. A8C1 and A8C2 present identical binding profiles (Figure 3.2), but cross-reactivity is more prone to occur during the synthesis of A8C1, as C1 component has both amine and carboxyl moieties (§Table 2.2).

Ligands A6A4 and A8A3 were both selected to proceed with the studies, as they present similar binding behaviour and have significant structural differences. However, it is relevant to mention that A4, A6, and A8 components may react in different ways in the triazine reaction. Tyramine (A4) is able to react with cyanuric chloride both through its amine and hydroxyl functionalities, as they are both strong nucleophiles. Nevertheless, if the synthesis is performed at neutral pH, the hydroxyl group will be almost unreactive [26]. Tryptamine (A6) is composed of an indole ring and an aliphatic primary amine, and histamine (A8) comprises an imidazole group and also an aliphatic primary amine (§Table 2.3). The amine groups of the indole and the imidazole heterocycles are able to react as nucleophiles, as well as the aliphatic primary amines. However, the reaction will favour the incorporation of the primary amine in cases of significant steric hindrance for the ring nitrogen. Moreover, it is important to guarantee that the reaction is carried out at alkaline conditions, above the p*K*a values of the amine groups of the heterocycles [27, 28].

A3B6 was selected as the lead Petasis-Ugi ligand, due to its superior performance in terms of binding capacity and selectivity when compared to A2B6 (Figure 3.2).

Best binding conditions were determined using a mixture of SW6-P, TW6-P, and YW6-P (1:1:1) (phosphopeptides mixture). Binding optimization was performed using a mixture of peptides, rather than the individual peptides, because it is important to understand the behaviour of the phosphopeptide targets when they are screened in more complex mixtures. Ligands were screened in parallel against the phosphopeptides mixture and a mixture of SW6, TW6, and YW6 (non-phosphopeptides mixture), in order to determine the selectivity of the ligands towards the phosphorylated peptides. In this case, phosphopeptides mass fractions were calculated instead of phosphopeptides mole fractions, because it is not possible to distinguish which of the peptides in each mixture are binding to the ligands. It is only possible to determine the total amount (in mg) of peptides bound in each screening. The phosphopeptides mass fraction (%) as utilized in this work is defined in Equation 3.1.

Equation 3.1 – Phosphopeptides mass fraction (%).

$$Phosphopeptides\ Mass\ Fraction\ (\%) = \frac{\frac{mg\ phosphopeptides}{g\ support}}{\frac{mg\ phosphopeptides}{g\ support} + \frac{mg\ non-phosphopeptides}{g\ support}} x100$$

Ligands A6A4, A8A3, A8C2, and A3B6 were screened against the phosphopeptides and the non-phosphopeptides mixtures at different pH (3 to 11). Higher binding capacities (mg phosphopeptide mixture/ g ligand-functionalized support) were obtained at acidic pH, with the highest value at pH 4 (Figure 3.5A). Ligands presented higher selectivity for phosphopeptides at pH 7 (Figure 3.5B), which might be related with the p*K*a values of pSer (pKa_1=1.73; pKa_2=6.73), pThr (pKa_1=1.70; pKa_2=6.71), and pTyr (pKa_1=1.26; pKa_2=6.24) (Figure 3.6). All phosphopeptides present two negative charges at pH values equal to or above 7, and one negative charge below it.

High binding capacities are combined with high enrichment efficiencies at pH 3 (Figure 3.5). At this pH, the carboxyl groups of the peptides are protonated and, therefore, their interference is minimized. The non-phosphorylated peptides do not present negative charges, while the phosphorylated peptides have one negative charge in the phosphate group (Figure 3.6). On the other hand, all ligands are positively charged at pH 3, being able to selectively bind the phosphopeptides through their negatively charged phosphate moiety. However, the lower binding capacities of the ligands at pH 3 relative to pH 4 might be associated to an incomplete deprotonation of the phosphate group of the peptides at more acidic pH. Tsao and co-workers have already reported that

carboxyl groups are only completely protonated at pH values below 2, whereas phosphate deprotonation is fully achieved at pH values above 3 [29].

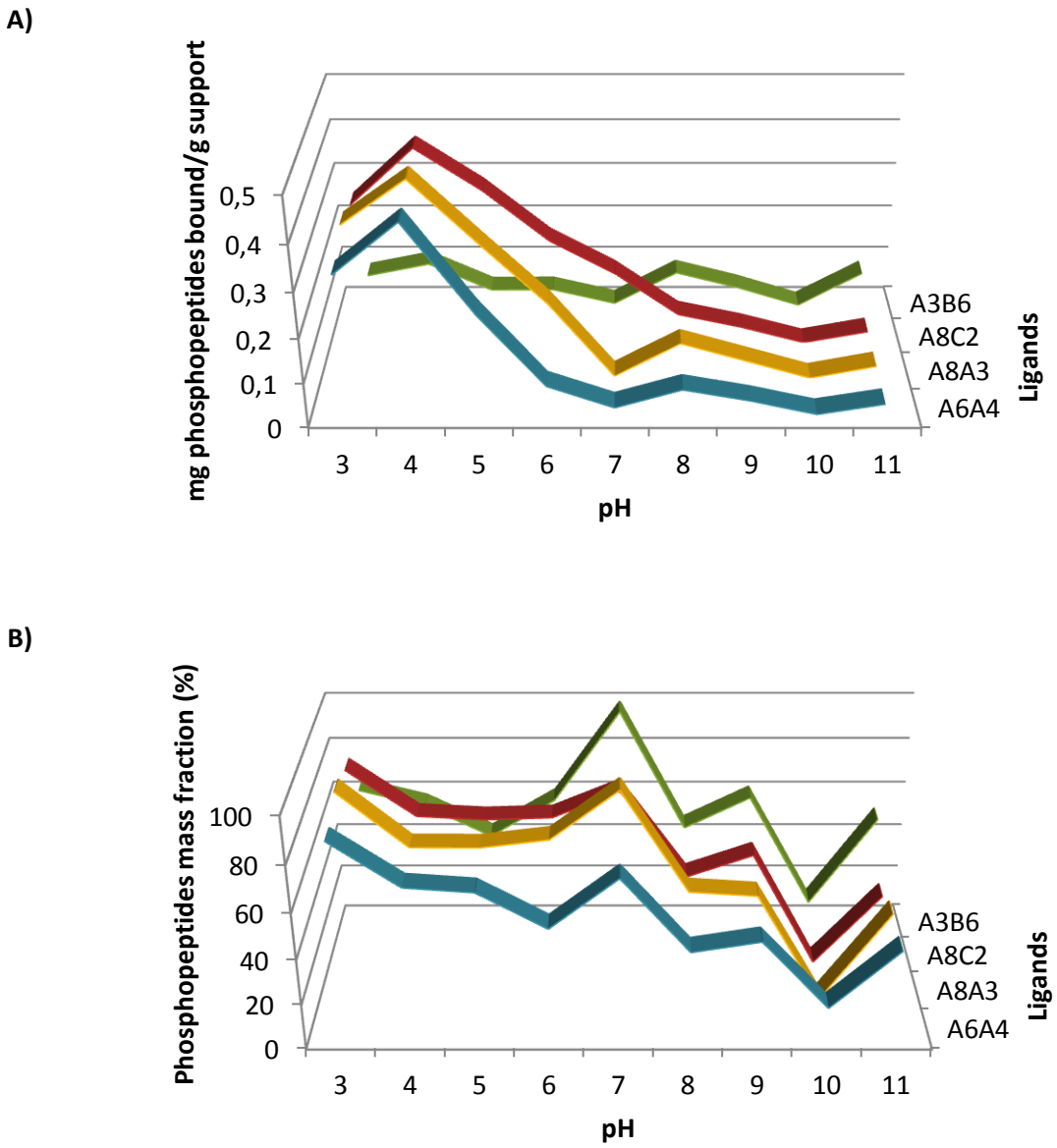


Figure 3.5 – Screening results of A6A4, A8A3, A8C2, and A3B6 ligands against a mixture of phosphopeptides at different pH. Screenings were performed by adding 0.25 mL of peptide mixture in binding buffer (0.5 mg/ mL) to each well of a 96-well filter plate containing 0.25 g of ligand-functionalized resin. The unbound peptides from flow-through and washes were collected in 96-well black microplates and the fluorescence intensity of the samples was measured in a microplate reader (λexc = 280 (20) nm – λem = 340 (35) nm filter). A) Amount of phosphopeptides mixture bound per g of ligand-functionalized support (N=2; stdev ≤ 0.02 mg/g). B) Phosphopeptides mass fraction (%) (N=2; stdev ≤ 19%). The binding buffers used were: 50 mM glycine pH 3; 50 mM acetate pH 4; 50 mM acetate pH 5; 50 mM acetate pH 6; 50 mM tris pH 7; 50 mM NH_4HCO_3 pH 8; 50 mM tris pH 9; 50 mM glycine pH 10; and 50 mM glycine pH 11.

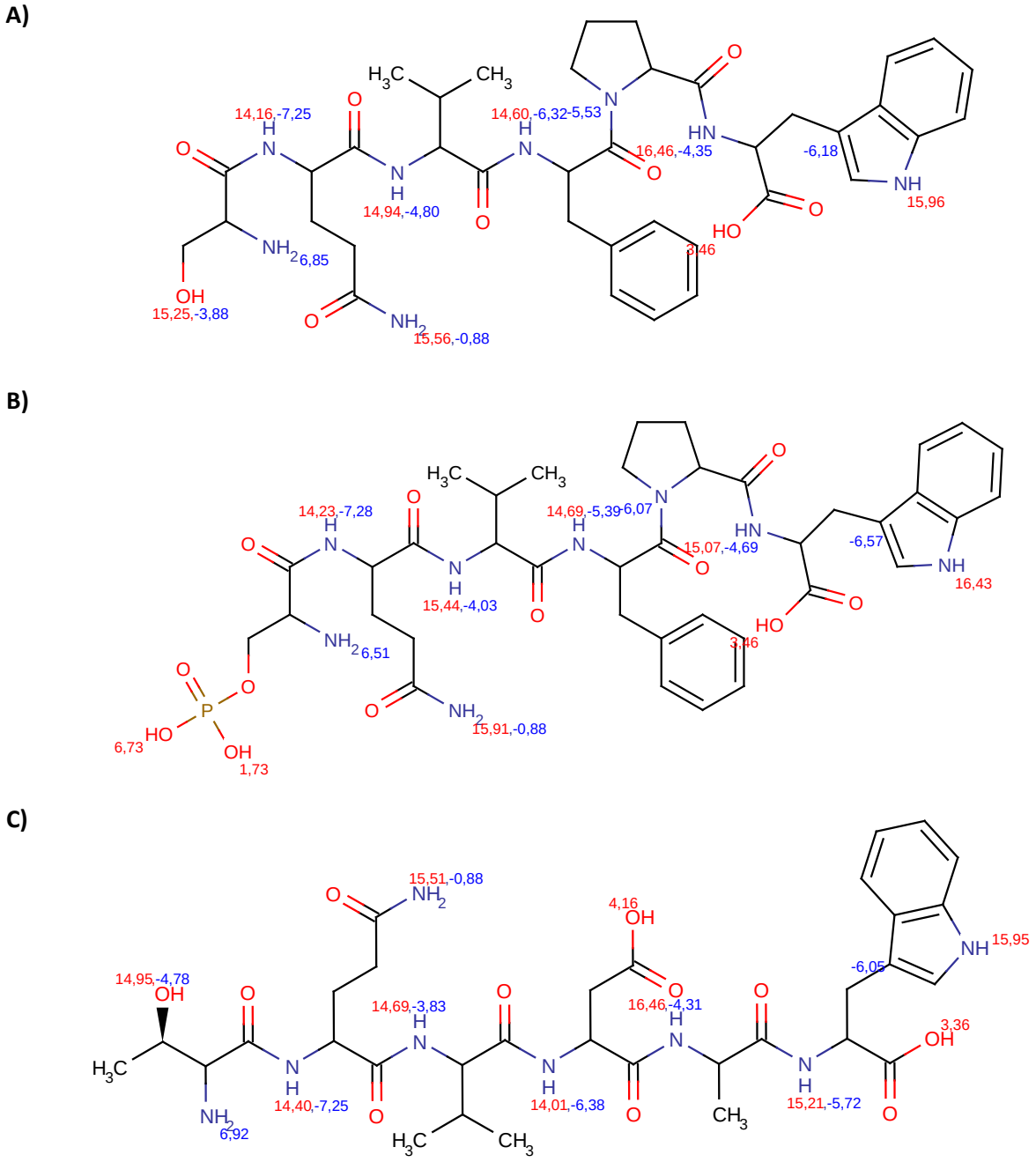
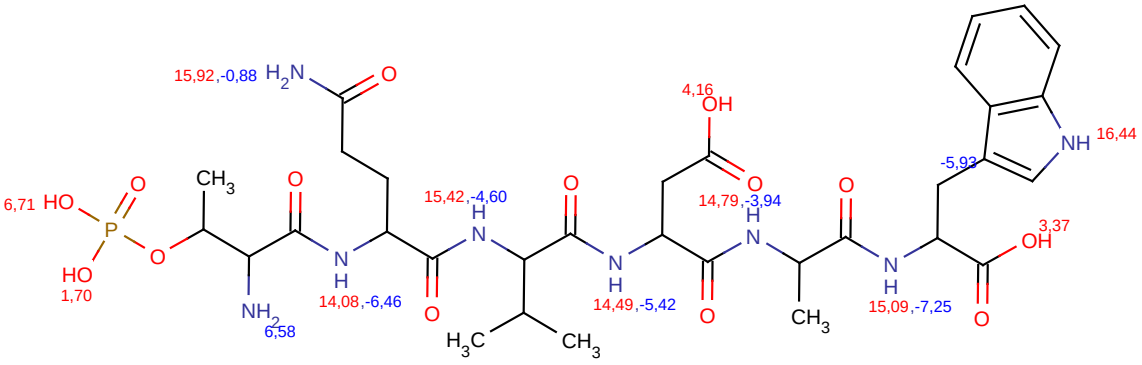
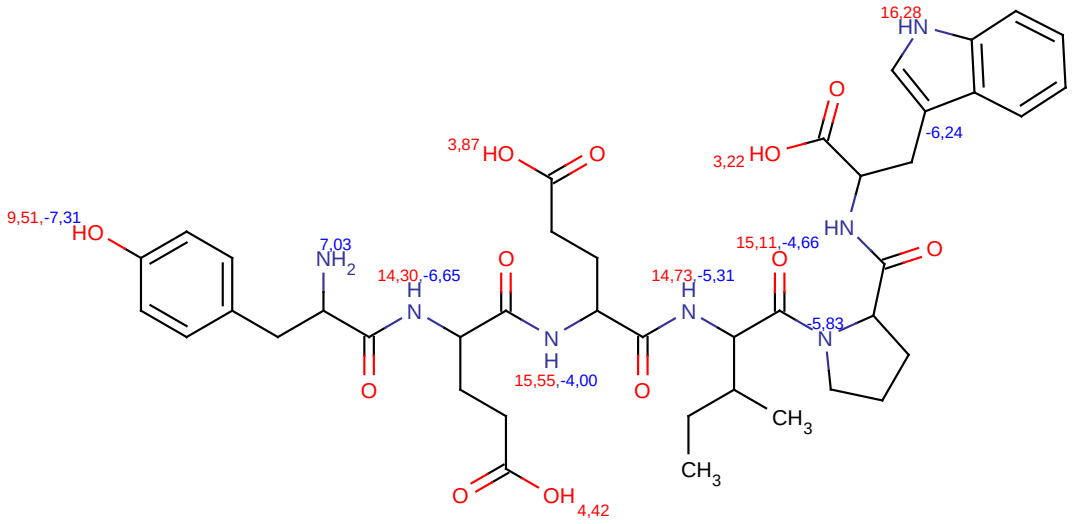


Figure 3.6 – p*Ka* values of functional groups from A) SW6; B) SW6-P; C) TW6; D) TW6-P; E) YW6; and F) YW6-P. Values were determined using the p*Ka* Plugin from MarvinSketch (ChemAxon).

D)



E)



F)

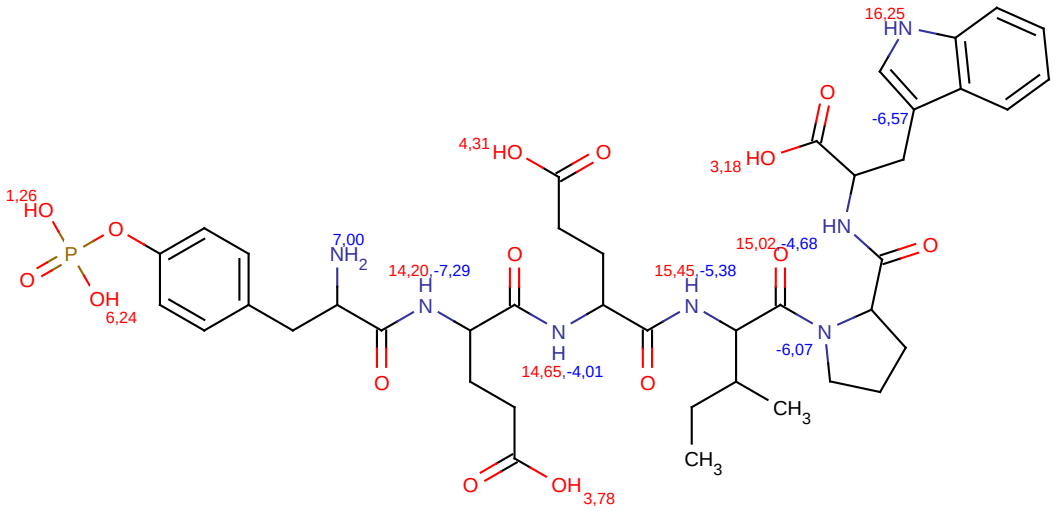


Figure 3.6 (cont.) – p*K*a values of functional groups from A) SW6; B) SW6-P; C) TW6; D) TW6-P; E) YW6; and F) YW6-P. Values were determined using the p*K*a Plugin from MarvinSketch (ChemAxon).

The influence of ionic strength in the binding profile of the peptides has also been investigated, since the addition of salt might decrease the non-specific binding of non-phosphorylated peptides. Therefore, both phosphopeptides and non-phosphopeptides mixtures were screened at pH 3 with increasing concentrations of NaCl (0, 25, 50, 100, 250, 500, 1000 mM). The presence of salt dramatically decreases the binding capacity of the ligands and their bias towards phosphopeptides, even at low concentrations. This effect has already been observed by Ndassa and co-workers, which reported that the addition of NaCl did not provide any improvement in phosphopeptide enrichment using immobilized IMAC [30].

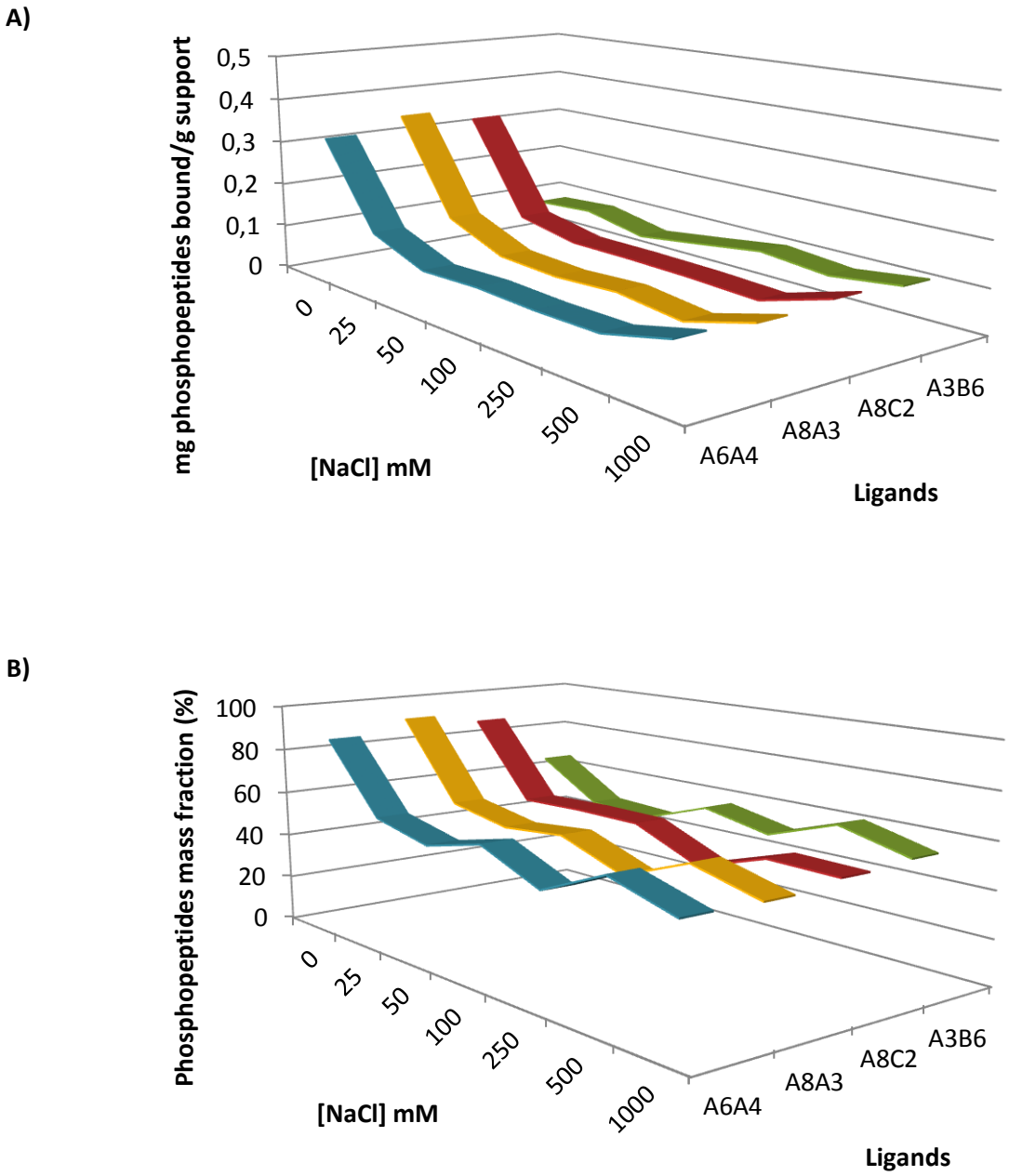


Figure 3.7 – Screening results of A6A4, A8A3, A8C2, and A3B6 ligands against a mixture of phosphopeptides at pH 3 with different salt concentrations. Screenings were performed by adding 0.25 mL of peptide mixture in binding buffer (0.5 mg/ mL) to each well of a 96-well filter plate containing 0.25 g of ligand-functionalized resin. The unbound peptides from flow-through and washes were collected in 96-well black microplates and the fluorescence intensity of the samples was measured in a microplate reader (λexc = 280 (20) nm – λem = 340 (35) nm filter). A) Amount of phosphopeptides mixture bound per g of ligand-functionalized support (N=2; stdev ≤ 0.03 mg/g). B) Phosphopeptides mass fraction (%) (N=2; stdev ≤ 17%).

Ligand A3B6 presented low binding capacities and selectivity at all pH values and salt concentrations, being not considered for subsequent studies. No significant differences were observed between triazine ligands; thus, the lead candidate was selected based on structural aspects. As previously stated, ligand A6A4 has two substituents which might originate side reactions and, consequently, this ligand was discarded. At this point, the two lead candidates (A8A3 and A8C2) were screened against a multi-phosphorylated peptide, in order to access their ability to bind molecules with a higher degree of phosphorylation. The multi-phosphorylated peptide has one pSer and two pThr, having the following sequence: Tyr-Ala-Gly-pSer-pThr-Asp-Glu-Asn-pThr-Asp-Ser-Glu-Trp (YW13-P). A shortened version of this peptide (without the Trp residue) has been reported to bind FHA domains of human polynucleotide kinase 3'-phosphatase (PNK) with high specificity and affinity (K_D = 190 nm) [31].

The first parameter changed during the screening of A8A3 and A8C2 against YW13-P and its non-phosphorylated version YW13 was the pH of the binding buffer. Ligands behaved in a significantly different manner in the presence of the multi-phosphorylated peptide, when compared to the screenings against the mixture of singly phosphorylated peptides. There is a minor decrease in the binding capacity of both ligands between pH 3 and pH 7. Ligand A8C2 binds approximately 90% of the multi-phosphorylated peptide at pH 7. Above this pH, the binding capacity of both ligands decreases considerably (Figure 3.8). One possible explanation is that the ligands are positively charged below pH 7, but are uncharged above it (Figure 3.9), while YW13-P is negatively charged in the entire pH range tested (

Figure **3.10**), as indicated by their p*K*a values. Therefore, electrostatic interactions between the ligands and YW13-P are more prone to occur below pH 7. However, it is important to keep in mind that the p*K*a values give an idea of the major microspecies of each molecule at a given pH, i.e. the most frequent protonation form. If the pH is near the p*K*a value of any particular functional group, the molecule will exist in several protonation forms. Also, the p*K*a values presented herein were calculated by software and may slightly differ from experimental p*K*a values. The imidazole group presents p*K*a values of 6.84 and 6.55 for ligands A8A3 and A8C2, respectively. Therefore, the ligands will also exist as their protonated forms (positively charged) at pH 7. On the other hand, each phosphate group presents two negative charges at this pH, and all carboxyl groups are negatively charged as well. This might be the reason why ligand A8C2 binds the highest amount of peptide at pH 7.

YW13-P and YW13 peptides are acidic peptides, having 5 carboxyl groups in their sequence. The fact that the ligands present phosphopeptide mass fractions around 50% below pH 7 indicates that the interaction is not occurring though the phosphate moiety in this case. At pH values of 7 or above,

the ligands are much more selective towards YW13-P, but at the expense of lower binding capacities (Figure 3.8).

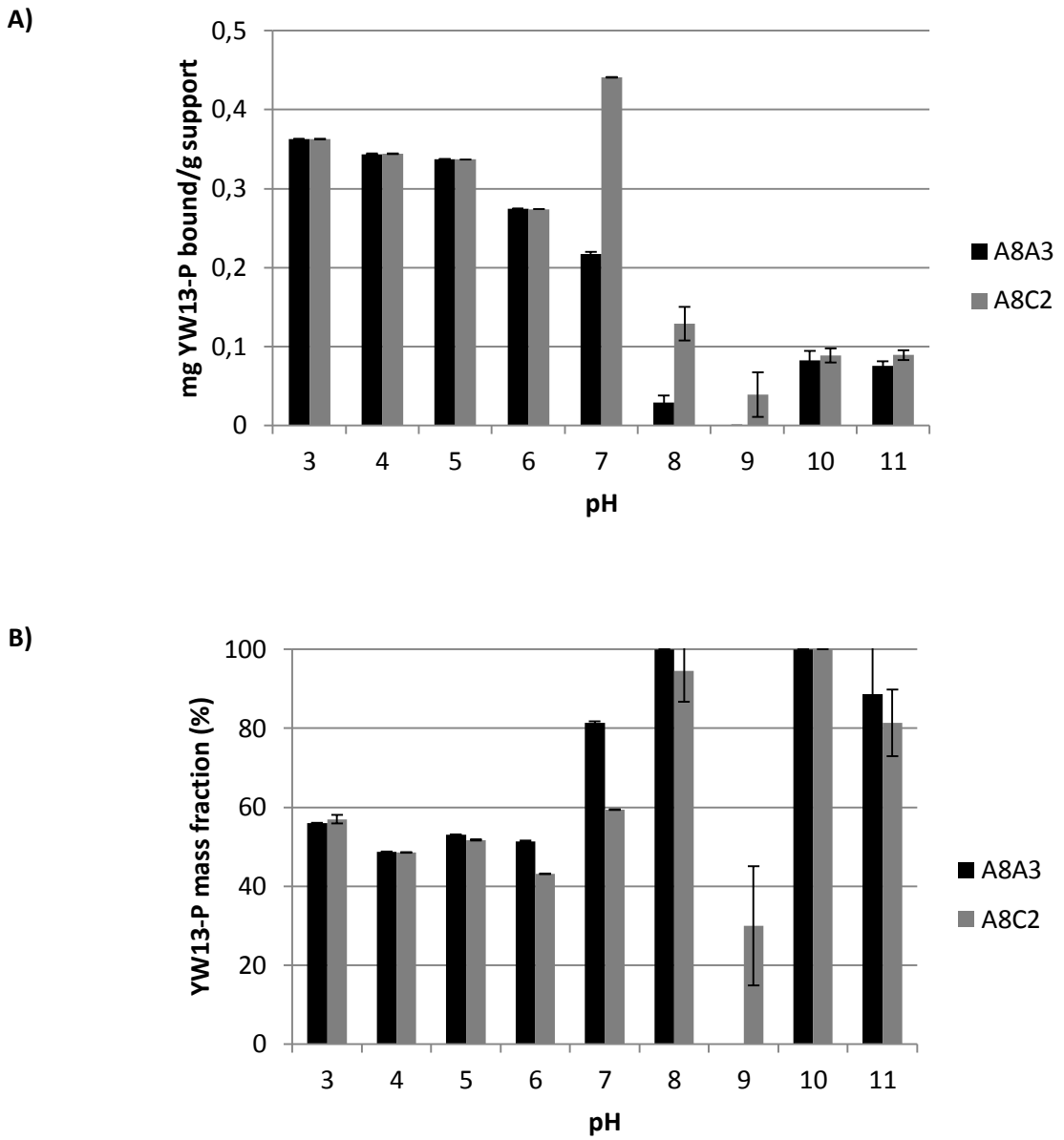
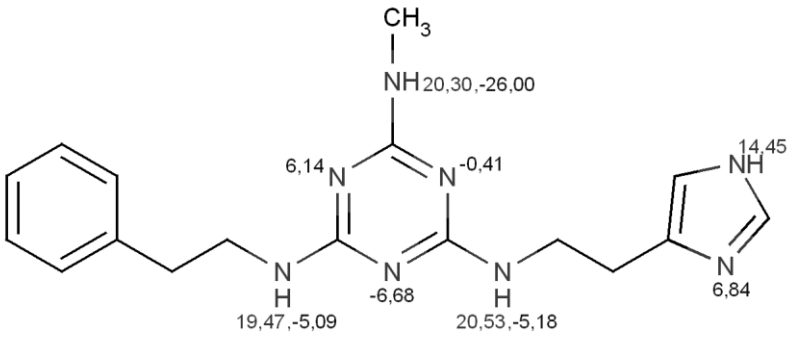


Figure 3.8 – Screening results of A8A3 and A8C2 against YW13-P at different pH. Screenings were performed by adding 0.25 mL of peptide in binding buffer (0.5 mg/ mL) to each well of a 96-well filter plate containing 0.25 g of ligand-functionalized resin. The unbound peptides from flow-through and washes were collected in 96-well black microplates and the fluorescence intensity of the samples was measured in a microplate reader (λexc = 280 (20) nm – λem = 340 (35) nm filter). A) Amount of YW13-P bound per g of ligand-functionalized support. B) YW13-P mass fraction (%). The binding buffers used were: 50 mM glycine pH 3; 50 mM acetate pH 4; 50 mM acetate pH 5; 50 mM acetate pH 6; 50 mM tris pH 7; 50 mM NH_4HCO_3 pH 8; 50 mM tris pH 9; 50 mM glycine pH 10; and 50 mM glycine pH 11. (N=2)

A)



B)

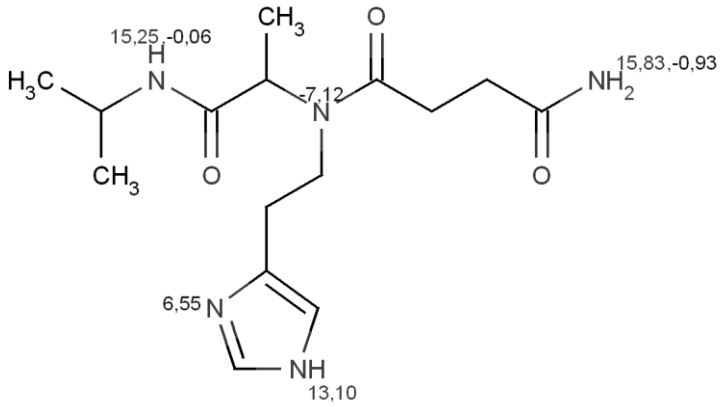


Figure 3.9 – p*K*a values of A) ligand A8A3, and B) ligand A8C2. Values were determined using the p*K*a Plugin from MarvinSketch (ChemAxon). A methyl group was included in the structure of the ligand in place of the agarose bead.

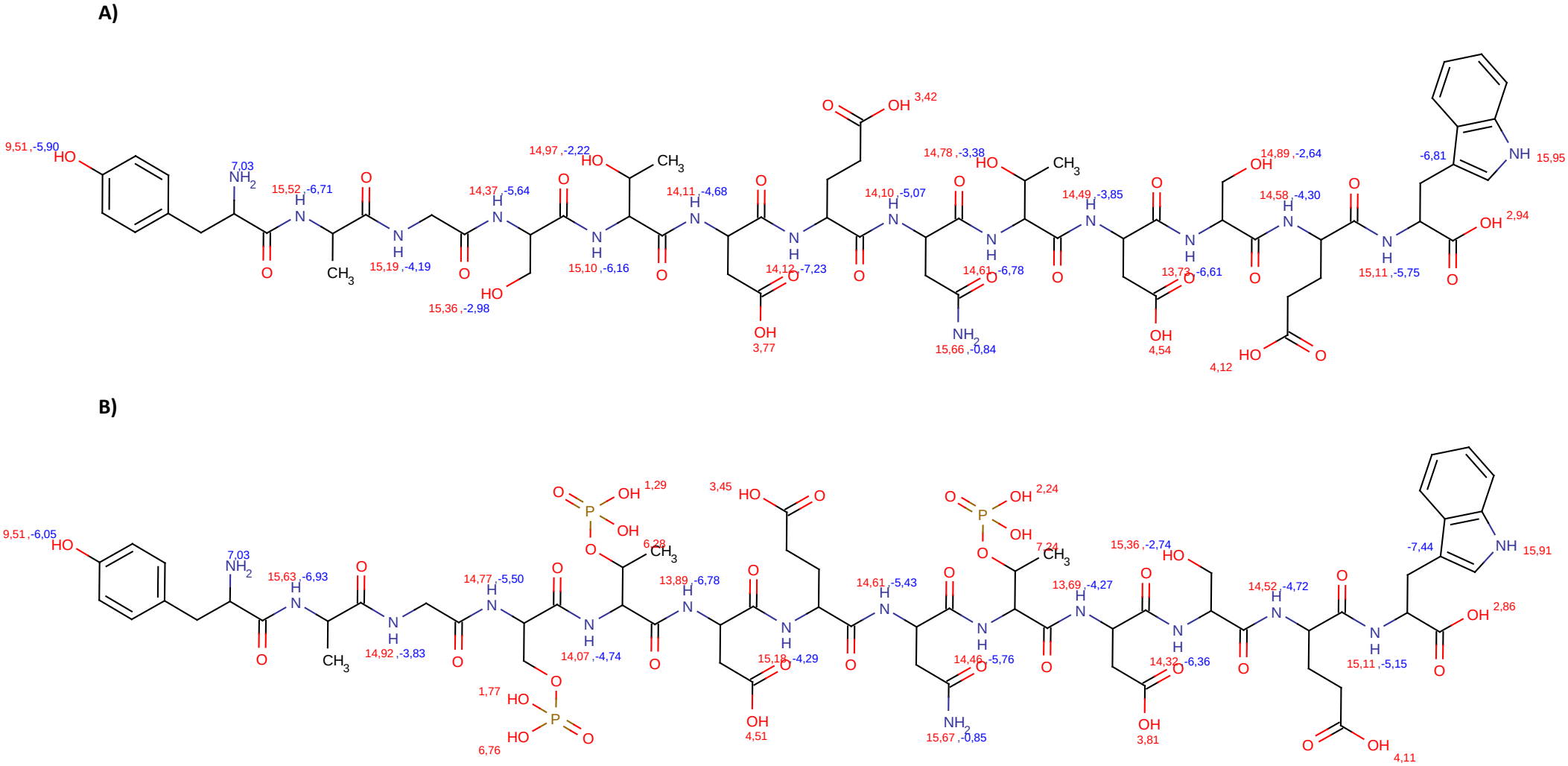


Figure 3.10 - p*Ka* values of A) YW13, and B) YW13-P. Values were determined using the p*Ka* Plugin from MarvinSketch (ChemAxon).

Ligands were also screened against YW13-P and YW13 at pH 3, but with different concentrations of salt, in order to compare their binding profile with the one from the mono-phosphorylated peptides mixture (Figure 3.11). The binding capacity of the ligands was not affected for low concentrations of salt, but the selectivity of the ligands significantly increased. Ligand A8A3 presented a phosphopeptide mass fraction of approximately 98% for a concentration of NaCl of 50 mM, while ligand A8C2 was almost 100% selective for a concentration of 100 mM. This is related to the fact that multi-phosphorylated peptides bind more strongly to the ligands, due to their enhanced negative charge, when compared to non- or mono-phosphorylated peptides. Therefore, the selectivity towards multi-phosphorylated peptides increases for low concentrations of salt. For concentrations of salt above 100 mM, there is a decrease in both the binding capacity and selectivity. This is in accordance to the results obtained by Nühse and co-workers, who used a two-dimensional methodology with strong anion exchange (SAX) chromatography prior to IMAC for the separation of phosphorylated peptides. They concluded that singly phosphorylated peptides eluted from the SAX resin at low concentrations of salt, while multi-phosphorylated peptides were retained, eluting at higher salt concentrations [32].

Loading two separate ligand-functionalized affinity columns, one with a sample with added salt, and one with the same sample without salt, would allow the selective enrichment of both mono- and multi-phosphorylated peptides from a complex sample.

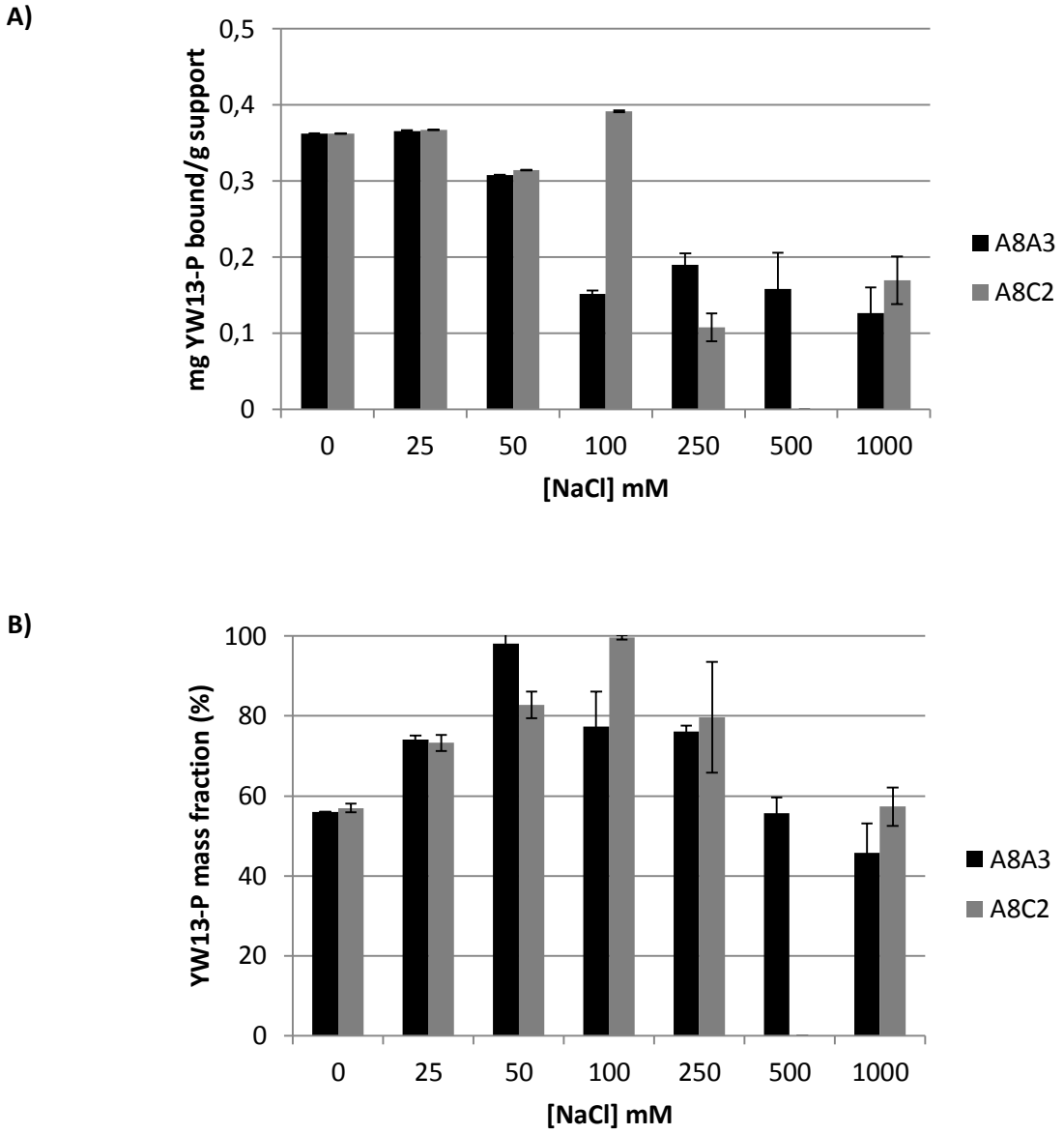


Figure 3.11 – Screening results of ligands A8A3 and A8C2 against YW13-P at pH 3 with different salt concentrations. Screenings were performed by adding 0.25 mL of peptide in binding buffer (0.5 mg/ mL) to each well of a 96-well filter plate containing 0.25 g of ligand-functionalized resin. The unbound peptides from flow-through and washes were collected in 96-well black microplates and the fluorescence intensity of the samples was measured in a microplate reader (λexc = 280 (20) nm – λem = 340 (35) nm filter). A) Amount of YW13-P bound per g of ligand-functionalized support; B) YW13-P mass fraction (%). (N=2)

3.3.3. Determination of binding capacity of negative controls

The triazine ligand A8A3 is synthesized on aminated agarose and the Ugi ligand A8C2 is synthesized on aldehyde-functionalized agarose. Therefore, agarose, agarose-NH_2, and agarose-CHO were used as negative controls to determine the non-specific binding of phosphorylated peptides to the solid support.

All individual peptides bind less than 20% to agarose. When agarose is screened against the mixture of mono-phosphorylated peptides, the value of non-specific binding corresponds to the sum of the values obtained for the individual peptides; thus, the effect is cumulative. Agarose-CHO presents an identical performance to agarose itself. The phosphopeptides binding profile to agarose-NH_2 is similar to the previous ones as well, with the exception of the multi-phosphorylated peptide YW13-P, which binds around 95%. At pH 3, amine groups will exist as NH_3^+, establishing stronger electrostatic interactions with the highly negatively charged multi-phosphorylated peptide. As a result, in the case of ligand A8A3, it is uncertain whether YW13-P is binding to unreacted amines or the ligand itself. However, as evidenced in chapter 2, triazine reaction presents satisfactory reaction yields, minimizing the interference of unreacted amines.

Table 3.5 – Non-specific binding of phosphopeptides (%) to agarose, agarose-CHO, and agarose-NH_2. Screenings were performed by adding 0.25 mL of peptide in 50 mM glycine pH 3 (0.5 mg/ mL) to each well of a 96-well filter plate containing 0.25 g solid support. The unbound peptides from flow-through and washes were collected in 96-well black microplates and the fluorescence intensity of the samples was measured in a microplate reader (λexc = 280 (20) nm – λem = 340 (35) nm filter). (N=2)

		Support		
		Agarose	Agarose-CHO	Agarose-NH_2
Peptides	SW6-P	2 ± 2	4 ± 5	3 ± 4
	TW6-P	2 ± 2	3 ± 4	0 ± 1
	YW6-P	19 ± 3	13 ± 5	20 ± 1
	Mixture of SW6-P, TW6-P, and YW6-P	22 ± 5	18 ± 4	26 ± 2
	YW13-P	1 ± 2	0 ± 0	95 ± 0.2

3.3.4. Optimization of elution conditions

Innumerable strategies have been used for the elution of phosphorylated proteins and peptides, depending on the enrichment method used. Peptides bound to ligand-functionalized agarose were eluted using different solutions, considering the following parameters: presence of an organic solvent, such as ACN or MeOH; use of acidic or alkaline solutions; type of acid added; and presence of a competitor molecule (H_3PO_4).

Organic solvents disrupt weak hydrophobic interactions between the peptides and the ligand. Lee and co-workers reported that including 40 to 50% ACN in the elution buffer increased the elution

efficiency of phosphorylated peptides from IMAC beads [33]. MeOH (Sol 2 and Sol 3 - Figure 3.12), for example, elutes both non- and mono-phosphorylated peptides, but not the multi-phosphorylated peptide YW13-P, which is strongly bound to the ligand through electrostatic interactions. If ACN is used instead of MeOH (Sol 1 - Figure 3.12), only mono-phosphorylated peptides are eluted, increasing the selectivity of the elution process. However, it should be taken into consideration that the elution efficiency is relative to the peptides bound to the ligands. As ligands are highly selective towards phosphorylated peptides, the actual percentage of non-phosphorylated peptides bound is rather low (< 15% for the mixture of SW6, TW6, and YW6; < 5% for YW13). Nevertheless, the elution process should be optimized in order to eliminate or minimize the co-elution of non-phosphorylated peptides.

Both acidic and alkaline conditions have been widely used for the recovery of phosphorylated peptides. Thingholm and co-workers introduced a strategy consisting in the sequential separation of mono- and multi-phosphorylated peptides from complex mixtures, combining both IMAC and TiO_2 chromatography. They have reported the elution of singly phosphorylated peptides under acidic conditions, whereas multiply phosphorylated peptides are subsequently recovered under basic conditions [34]. In other study, a basic eluent (NH_4OH) has been described to efficiently elute both mono- and multi-phosphorylated peptides from IMAC [35]. The elution efficiency of different bases has been reported for TiO_2 chromatography, and higher phosphopeptide recoveries were obtained with NH_4OH and $NH_4H_2PO_4$ eluents, when compared to NH_4HCO_3 [36].

In this study, a solution of 50% (v/v) ACN/H_2O + 0.1% NH_4OH pH 10.2 (Sol 4 - Figure 3.12) allows the elution of both mono- and multi-phosphorylated peptides from ligand A8A3, but only mono-phosphorylated peptides and a few non-phosphorylated peptides are eluted from ligand A8C2. At pH 10.2, both A8A3 and A8C2 ligands are uncharged and present logD values of 2.19 and -1.58, respectively. Therefore, ligand A8A3 is hydrophobic at pH 10.2, while ligand A8C2 is hydrophilic. This might be the reason why phosphorylated peptides elute more easily from ligand A8A3, since they mainly interact with the ligand through electrostatic interactions. The presence of ACN further enhances this effect, as it further disrupts weak hydrophobic interactions that might also exist between the ligand and the peptide.

The influence of weak acids upon elution was evaluated using H_3BO_3 and H_3PO_4. H_3BO_3 did not lead to the elution of neither phosphorylated nor non-phosphorylated peptides (Sol 7 - Figure 3.12). However, when a strong acid (TFA) was added to the solution both mono- and multi-phosphorylated peptides were eluted (Sol 8 - Figure 3.12). A solution comprising 50% (v/v) ACN/H_2O+ 0.1% TFA + 1% H_3BO_3 allowed for 94% recovery of YW13-P from ligand A8A3 (Sol 8 - Figure 3.12A). Thingholm and

co-workers used increasing concentrations of TFA and were able to elute mono-phosphorylated peptides for TFA concentrations between 0.2 and 1% and multi-phosphorylated peptides for TFA concentrations above 1% from IMAC adsorbents [34].

The use of 1% H_3PO_4 gave the best results, allowing the recovery of both mono- and multi-phosphorylated peptides from both affinity ligands (Sol 5 and Sol 6 - Figure 3.12). H_3PO_4 plays two roles, acting as an acid and as a competitor molecule. In this case, adding TFA to the solution did not improve the recovery of phosphorylated peptides (Sol 6 - Figure 3.12). A combination of ACN and H_3PO_4 has been reported to be extremely effective in the elution of both mono- and multi-phosphorylated peptides from IMAC [37].

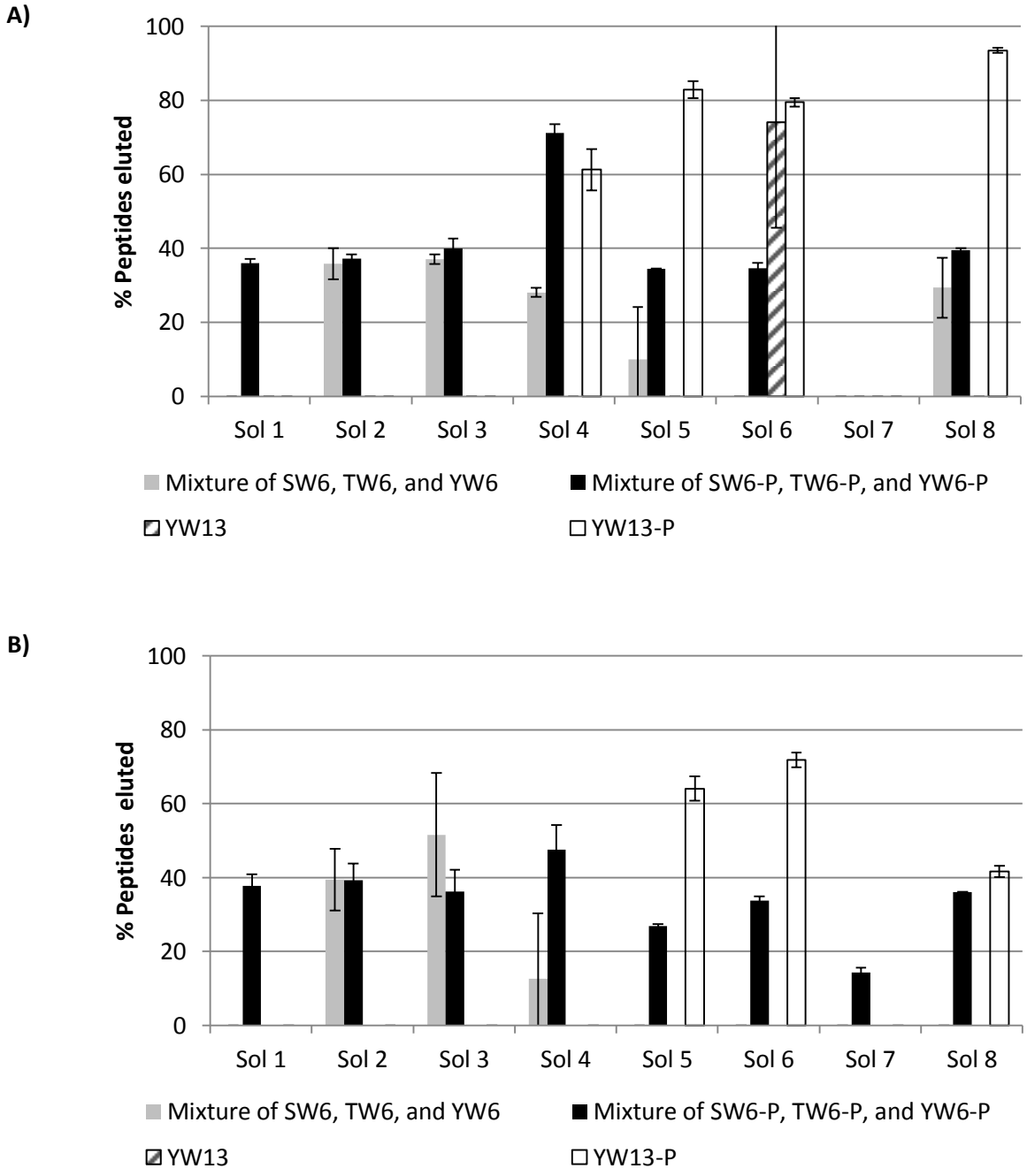


Figure 3.12 – Elution of phosphorylated and non-phosphorylated peptides from A) ligand A8A3 and B) ligand A8C2. After screening, peptides bound to ligand-functionalized supports were eluted by adding 0.25 mL of elution solution to each well of the 96-well filter plate. This step was repeated until no peptide was detected in the elution fractions. The eluted peptides were collected in 96-well black microplates and the fluorescence intensity of the samples was measured in a microplate reader (λexc = 280 (20) nm – λem = 340 (35) nm filter). The elution solutions were: Sol 1 – 50% (v/v) ACN/H_2O + 1% acetic acid (pH 3); Sol 2 – 50% (v/v) MeOH/H_2O + 1% acetic acid (pH 3); Sol 3 – MeOH (pH 6.3); Sol 4 – 50% (v/v) ACN/H_2O + 0.1% NH_4OH (pH 10.2); Sol 5 – 50% (v/v) ACN/H_2O + 1% H_3PO_4 (pH 1.6); Sol 6 – 50% (v/v) ACN/H_2O + 0.1% TFA + 1% H_3PO_4 (pH 1.3); Sol 7 - 50% (v/v) ACN/H_2O + 1% H_3BO_3 (pH 4.9); Sol 8 - 50% (v/v) ACN/H_2O + 0.1% TFA + 1% H_3BO_3 (pH 1.5). (N=2)

3.3.5. Partition equilibrium experiments

A simple bimolecular interaction between a target peptide and a solid-phase ligand is depicted in Figure 3.13, where K_A is the association equilibrium constant and K_D is the dissociation equilibrium constant. K_A is defined as in Equation 3.2, and K_D is $1/K_A$.

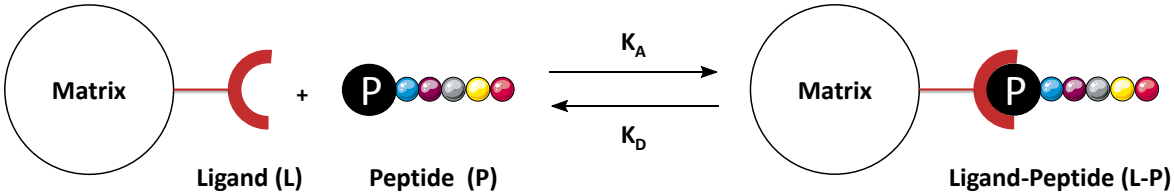


Figure 3.13 – Schematic representation of the interaction between an immobilized affinity ligand and a target peptide.

Equation 3.2 – Association equilibrium constant. [L] is the concentration of the ligand, [P] is the concentration of the peptide, and [L-P] is the concentration of the ligand-peptide complex.

$$K_A = \frac{[L-P]}{[L][P]}$$

Herein, the binding constants were determined by partition equilibrium studies, but other approaches may be used depending on the suitability of each method for a particular system [38]. The experiments consisted of incubating a known amount of partitioning solute (SW6, SW6-P, TW6, TW6-P, YW6, YW6-P, YW13, YW13-P) with a fixed amount of affinity matrix (A8A3 and A8C2 functionalized agarose) at constant temperature, until the mixture reach a chemical equilibrium. The adsorption of all phosphorylated and non-phosphorylated peptides followed the Langmuir adsorption model for both ligands (Figure 3.14 and Figure 3.15). The model relies on the following main assumptions: (i) each molecule occupies one adsorption site, forming a monolayer at the surface; (ii) every adsorption site is energetically equivalent; (iii) molecules adsorb independently of one another; and (iv) molecules do not interact with each other [39]. The Langmuir adsorption isotherm is represented by Equation 3.3.

Equation 3.3 – Langmuir adsorption isotherm. q is the binding capacity (mg peptide bound/g support); Q_{max} is the maximum binding capacity (mg peptide bound/g support); C_{eq} is the concentration of the peptide in solution at the equilibrium (mg/mL); and K_A is the association equilibrium constant (mL/mg).

$$q = \frac{Q_{max}\ K_A\ C_{eq}}{1 + K_A\ C_{eq}}$$

The K_A (M^{-1}) and Q_{max} (mg peptide bound/g support) parameters were determined and are recorded in Table 3.6 and Table 3.7.

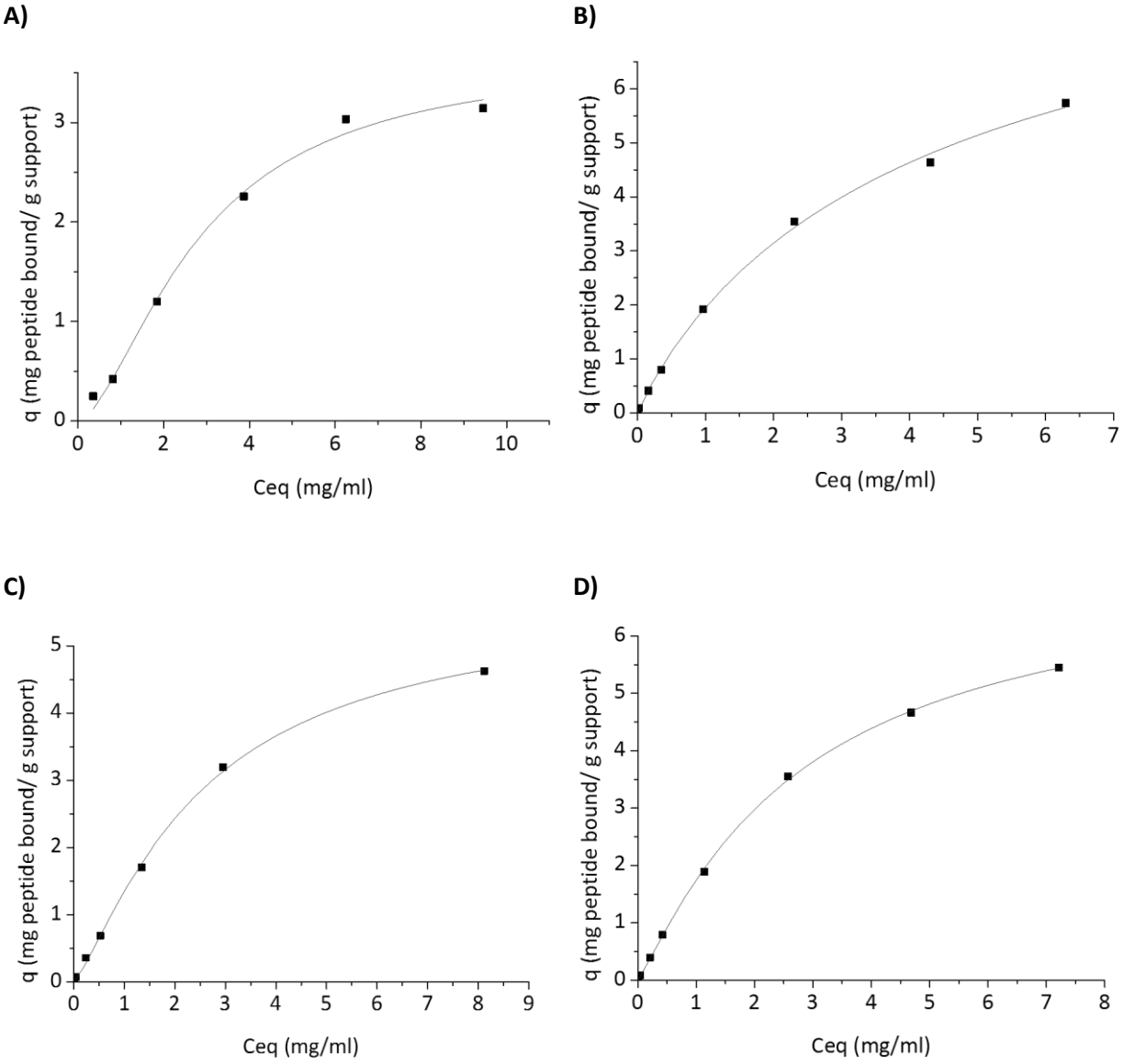


Figure 3.14 – Adsorption isotherms of A) SW6; B) SW6-P; C) TW6; D) TW6-P; E) YW6; F) YW6-P; G) YW13; and H) YW13-P for ligand A8A3. Peptides in different concentrations (0-10 mg/mL) were incubated overnight with 0.25 g of A8A3-functionalized agarose at 30°C. The buffers used were 50 mM Gly pH 3 for SW6, SW6-P, TW6, TW6-P, YW6, YW6-P, and 50 mM Gly pH 3 50 mM NaCl for YW13 and YW13-P. The unbound peptide from flow-through was collected and the fluorescence intensity of the samples was measured in a microplate reader (λexc = 280 (20) nm – λem = 340 (35) nm filter). The Langmuir equation was successfully fitted to data by nonlinear regression using OriginPro 8.5.1 SR2. (N=2)

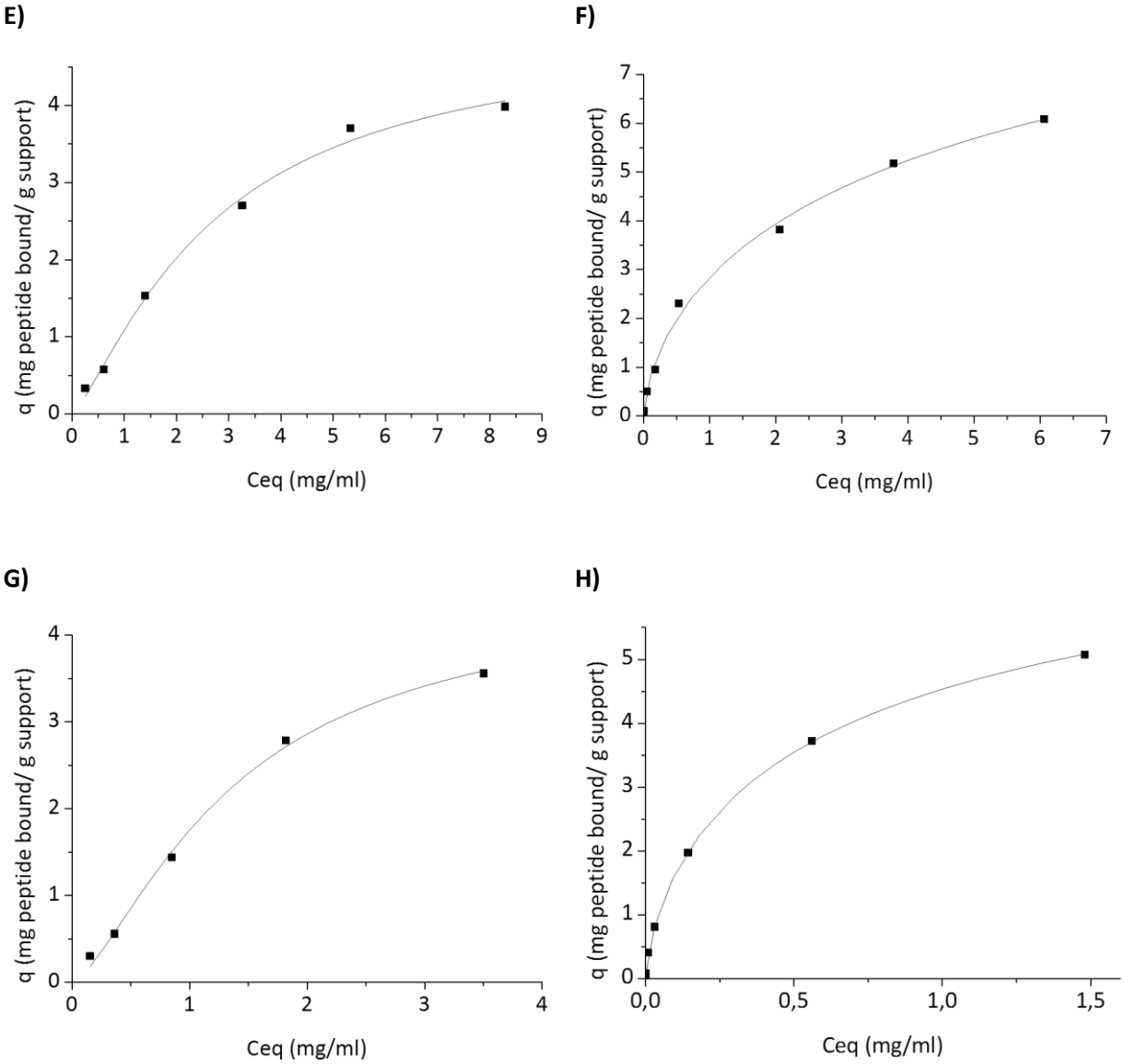


Figure 3.14 (cont.) – Adsorption isotherms of A) SW6; B) SW6-P; C) TW6; D) TW6-P; E) YW6; F) YW6-P; G) YW13; and H) YW13-P for ligand A8A3. Peptides in different concentrations (0-10 mg/mL) were incubated overnight with 0.25 g of A8A3-functionalized agarose at 30°C. The buffers used were 50 mM Gly pH 3 for SW6, SW6-P, TW6, TW6-P, YW6, YW6-P, and 50 mM Gly pH 3 50 mM NaCl for YW13 and YW13-P. The unbound peptide from flow-through was collected and the fluorescence intensity of the samples was measured in a microplate reader (λexc = 280 (20) nm – λem = 340 (35) nm filter). The Langmuir equation was successfully fitted to data by nonlinear regression using OriginPro 8.5.1 SR2. (N=2)

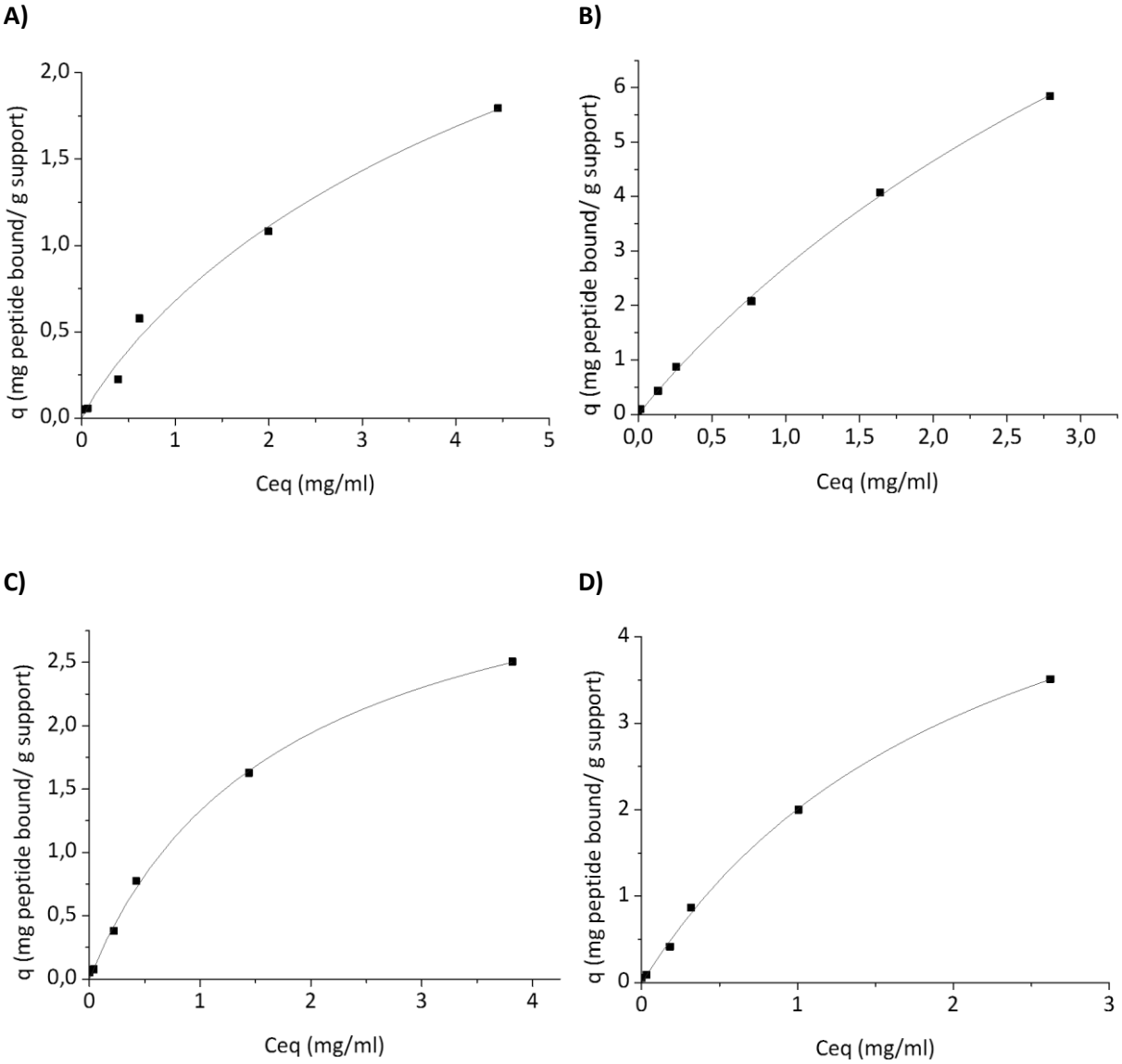


Figure 3.15 – Adsorption isotherms of A) SW6; B) SW6-P; C) TW6; D) TW6-P; E) YW6; F) YW6-P; G) YW13; and H) YW13-P for ligand A8C2. Peptides in different concentrations (0-10 mg/mL) were incubated overnight with 0.25 g of A8C2-functionalized agarose at 30°C. The buffers used were 50 mM Gly pH 3 for SW6, SW6-P, TW6, TW6-P, YW6, YW6-P, and 50 mM Gly pH 3 100 mM NaCl for YW13 and YW13-P. The unbound peptide from flow-through was collected and the fluorescence intensity of the samples was measured in a microplate reader (λexc = 280 (20) nm – λem = 340 (35) nm filter). The Langmuir equation was successfully fitted to data by nonlinear regression using OriginPro 8.5.1 SR2. (N=2)

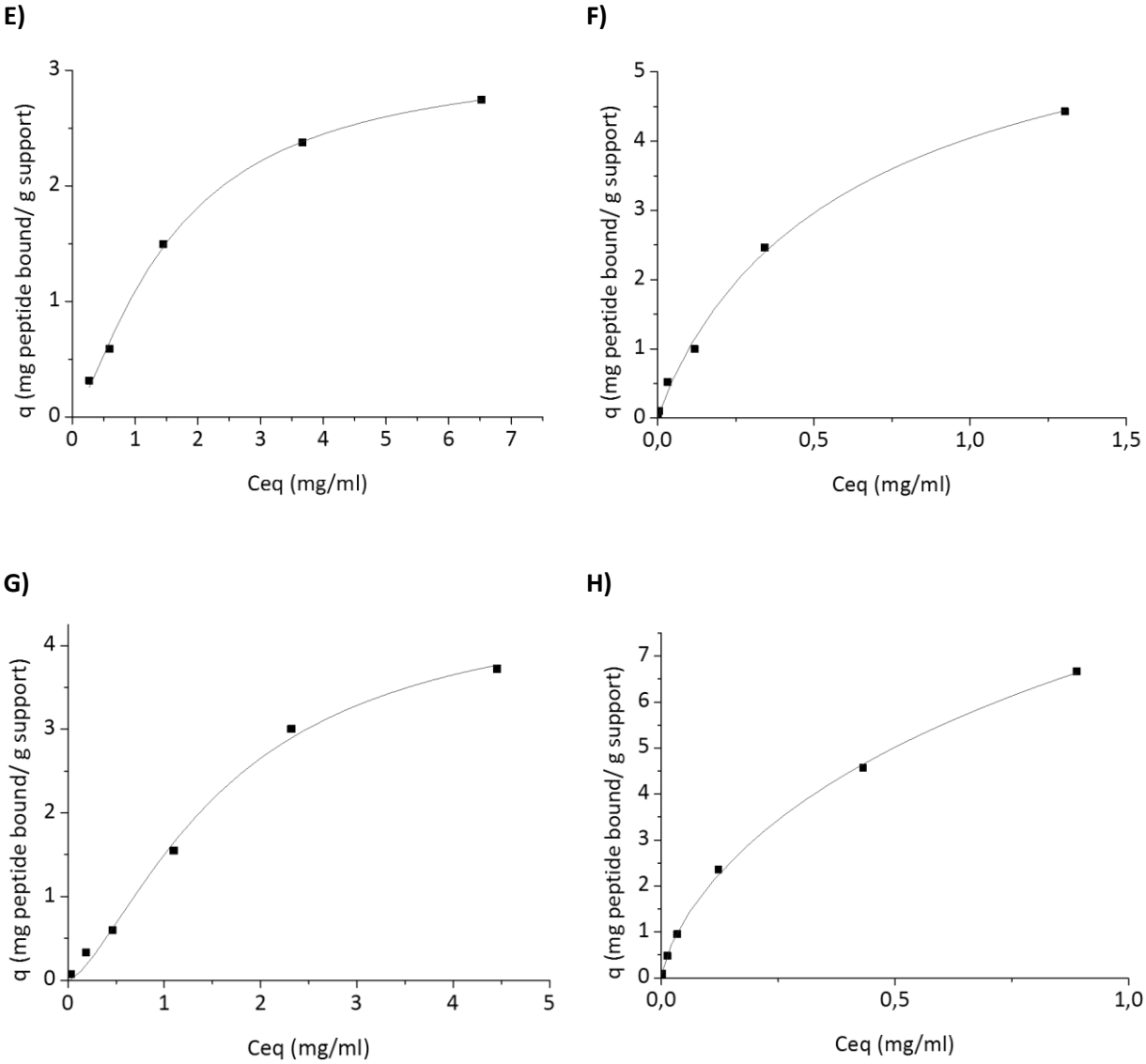


Figure 3.15 (cont.) – Adsorption isotherms of A) SW6; B) SW6-P; C) TW6; D) TW6-P; E) YW6; F) YW6-P; G) YW13; and H) YW13-P for ligand A8C2. Peptides in different concentrations (0-10 mg/mL) were incubated overnight with 0.25 g of A8C2-functionalized agarose at 30°C. The buffers used were 50 mM Gly pH 3 for SW6, SW6-P, TW6, TW6-P, YW6, YW6-P, and 50 mM Gly pH 3 100 mM NaCl for YW13 and YW13-P. The unbound peptide from flow-through was collected and the fluorescence intensity of the samples was measured in a microplate reader (λexc = 280 (20) nm – λem = 340 (35) nm filter). The Langmuir equation was successfully fitted to data by nonlinear regression using OriginPro 8.5.1 SR2. (N=2)

Table 3.6 – Q_{max} (mg peptide bound/ g support) and K_A (M^{-1}) parameters for ligand A8A3. Data was fitted to Langmuir equation on OriginPro 8.5.1 SR2. R^2 is the correlation factor. (N=2)

	SW6	SW6-P	TW6	TW6-P	YW6	YW6-P	YW13	YW13-P
Q_{max} (mg/g)	3.68 ±0.33	9.82 ±1.15	5.69 ±0.26	7.29 ±0.30	5.02 ±0.58	13.07 ±4.26	4.50 ±0.51	8.34 ±0.26
K_A (x10^2 M^{-1})	1.41 ±0.25	2.10 ±0.42	2.24 ±0.16	2.53 ±0.15	2.32 ±0.38	2.55 ±1.14	9.42 ±2.24	20.4 ±1.40
R^2	0.990	0.998	0.998	0.999	0.992	0.996	0.993	0.999

Table 3.7 – Q_{max} (mg peptide bound/ g support) and K_A (M^{-1}) parameters for ligand A8C2. Data was fitted to Langmuir equation on OriginPro 8.5.1 SR2. R^2 is the correlation factor. (N=2)

	SW6	SW6-P	TW6	TW6-P	YW6	YW6-P	YW13	YW13-P
Q_{max} (mg/g)	4.36 ±3.98	18.27 ±5.44	3.73 ±0.47	6.44 ±1.24	3.14 ±0.18	6.83 ±1.42	4.61 ±0.52	19.32 ±4.93
K_A (x10^2 M^{-1})	1.41 ±1.59	1.47 ±0.53	3.98 ±0.92	3.62 ±1.14	4.45 ±0.54	13.32 ±7.04	7.14 ±1.49	9.73 ±3.76
R^2	0.983	0.999	0.997	0.998	0.996	0.995	0.993	0.999

The low values of the affinity constants ($K_A < 10^3$ M^{-1}) indicate that the adsorption of the peptides to the ligands is mainly governed by weak interactions [40]. Both phosphorylated and non-phosphorylated peptides present binding affinities within the same range, which means the presence of the phosphate group does not significantly increases the stability of the peptide-ligand complex. It is important, however, to properly distinguish between the concepts of affinity and specificity, although they are both involved in molecular recognition [41]. Affinity may be defined as the structural and energetic recognition between two molecules [42]. It is a quantifiable property and is characterized by thermodynamical parameters, such as heat capacity and free enthalpy of dissociation at equilibrium [43, 44]. It determines the stability of the complex between the two molecules, which interact by non-covalent bonding. Specificity is the ability of one molecule to discriminate between its target and other competing molecules [44, 45]. It takes into consideration that the binding between the two molecules occurs in highly heterogeneous environments. It accounts not only for the existence of competing species with different relative concentrations, but also for the fact that the target molecule might bind outside its specific site or in different

orientations [41]. Specificity is a much more complex concept than affinity, being more accurately described by statistical physics [41, 43, 44]. High affinity systems do not necessarily present high specificity and vice-versa [41, 43].

High affinity is usually characterized by K_A values between 10^3 and 10^9 M^{-1}. However, these values have been determined for systems comprising the interaction between two macromolecules or a small molecule and a macromolecule [40]. In this work, both ligands and target peptides are small molecules, which are usually less flexible and have smaller surface areas. Larger molecules such as proteins have high structural and functional complexity. Their adsorption is governed by numerous factors, including structural re-arrangements; changing surface affinities upon binding; cooperative effects; overshooting adsorption kinetics; and aggregation. Although small molecules present higher diffusion rates, larger molecules usually bind more strongly to their target due to the larger contact area [46].

Eriksson and co-workers studied the affinity interaction between several phosphorylated peptides and a TiO_2 surface. They concluded that the K_A increases 2-fold per phosphorylation site, i.e. the affinity increases with the degree of phosphorylation, although it might be affected by other factors, such as the primary sequence of the peptide, the position of the phosphate groups, the existence of other charged amino acids, etc. The K_A values between the phosphopeptides and TiO_2 were in the order of 10^6 M^{-1} [47]. However, TiO_2 interacts with the phosphopeptides by metal coordination, through the formation of multi-dentate bonds. The simultaneous coordination to several metal ions may increase the binding affinity by up to a factor of 1000 [48].

Regarding the maximum binding capacities, all phosphorylated peptides presented higher Q_{max} than the non-phosphorylated peptides. As the partition equilibrium experiments were performed simultaneously for all peptides and used the same batch of synthesized ligand, it is reasonable to assume a constant density of ligand. Therefore, high Q_{max} values for phosphorylated peptides (even though having low binding affinities) might be a consequence of higher specificity of the ligands for phosphorylated peptides.

In a general chromatographic perspective, the capacity factor (k'), which determines the retention of a particular target, is related to K_A and Q_{max} by Equation 3.4. This equation is only valid when $K_A C_{eq} < 1$ (C_{eq} is the concentration of the target in solution at the equilibrium).

Equation 3.4 – Mathematical definition of capacity factor (k'). C is a constant characteristic from the support, and usually ranges between 0.5 and 1. Q_{max} is the maximum binding capacity and K_A is the affinity constant.

$$k' = C\ Q_{max}\ K_A$$

Weak affinity interaction chromatography relies on the dynamic interaction between ligand and target (rather than adsorption/ desorption phenomenon), being characterized by low K_A values and requiring high Q_{max}. This mode of affinity chromatography is thought to rely on repeated binding events or simultaneous weak binding events, presenting high efficiency and improved performance in terms of resolving power [49, 50].

3.4. Conclusions

Synthetic affinity ligands based on different chemistries have been successfully used to enrich phosphorylated peptides recognized by different PBDs, namely BRCT, FHA, and SH2 domains. Best ligands were based on triazine (A8A3) and Ugi (A8C2) reactions, and comprise an imidazole group in their structures. These ligands can be considered as multimodal ligands, as they have more than one mode of interaction with their targets. Particularly, the presence of heterocyclic groups, such as the imidazole group, provides them with the potential to interact both by electrostatic and hydrophobic interactions [51].

Best binding conditions combined with higher specificity were found to be pH 3 for a mixture of mono-phosphorylated peptides, and pH 7 for a multi-phosphorylated peptide. However, the multi-phosphorylated peptide was also successfully enriched at pH 3 when low concentrations of NaCl were added (up to 100 mM). A solution of 50% (v/v) ACN/H_2O + 1% H_3PO_4 (pH 1.6) allowed the recovery of both mono- and multi-phosphorylated peptides from both ligands, but the use of 50% (v/v) ACN/H_2O + 0.1% NH_4OH (pH 10.2) was effective as well.

All peptides presented low binding affinities ($K_A \approx 10^2$-10^3 M^{-1}) towards both ligands, but phosphorylated peptides presented higher Q_{max} values. This behaviour is typical from weak affinity interactions, which provide means for improved kinetics, high-resolution and rapid throughput [52, 53].

A full study by NMR, X-ray crystallography, or even molecular modelling is needed to better understand the interactions occurring between the ligands and their phosphorylated targets. Nonetheless, these tailored small molecules are promising alternatives to unspecific methods, such as IMAC and TiO_2 chromatography, contributing for the development of methodologies based on molecular recognition, which may have great impact not only in phosphopeptide enrichment at laboratory scale, but also as novel diagnostic tools.

3.5. Bibliography

1. Coffman, J.L., J.F. Kramarczyk, and B.D. Kelley, *High-throughput screening of chromatographic separations: I. Method development and column modeling.* Biotechnology and Bioengineering, 2008. **100**(4): p. 605-618.
2. Rege, K., et al., *High-throughput process development for recombinant protein purification.* Biotechnology and Bioengineering, 2006. **93**(4): p. 618-630.
3. Saraswat, L.D., et al., *Affinity ligand selection from a library of small molecules: Assay development, screening, and application.* Biotechnology Progress, 2005. **21**(1): p. 300-308.
4. Kramarczyk, J.F., B.D. Kelley, and J.L. Coffman, *High-throughput screening of chromatographic separations: II. Hydrophobic interaction.* Biotechnology and Bioengineering, 2008. **100**(4): p. 707-720.
5. Kelley, B.D., et al., *High-throughput screening of chromatographic separations: IV. Ion-exchange.* Biotechnology and Bioengineering, 2008. **100**(5): p. 950-963.
6. Curtis, R. and L. Lue, *A molecular approach to bioseparations: protein–protein and protein–salt interactions.* Chemical Engineering Science, 2006. **61**(3): p. 907-923.
7. Hofmeister, F., *On the understanding of the effects of salts.* Arch. Exp. Pathol. Pharmakol.(Leipzig), 1888. **24**: p. 247-260.
8. Broering, J.M. and A.S. Bommarius, *Evaluation of Hofmeister effects on the kinetic stability of proteins.* The Journal of Physical Chemistry B, 2005. **109**(43): p. 20612-20619.
9. Zhang, Y. and P.S. Cremer, *Interactions between macromolecules and ions: the Hofmeister series.* Current Opinion in Chemical Biology, 2006. **10**(6): p. 658-663.
10. Zangi, R., *Can salting-in/salting-out ions be classified as chaotropes/kosmotropes?* The Journal of Physical Chemistry B, 2009. **114**(1): p. 643-650.
11. Firer, M., *Efficient elution of functional proteins in affinity chromatography.* Journal of Biochemical and Biophysical Methods, 2001. **49**(1): p. 433-442.
12. Zhou, S., et al., *SH2 domains recognize specific phosphopeptide sequences.* Cell, 1993. **72**(5): p. 767-778.
13. Durocher, D., et al., *The molecular basis of FHA domain: phosphopeptide binding specificity and implications for phospho-dependent signaling mechanisms.* Molecular Cell, 2000. **6**(5): p. 1169-1182.
14. Csizmadia, F., et al., *Prediction of distribution coefficient from structure. 1. Estimation method.* Journal of Pharmaceutical Sciences, 1997. **86**(7): p. 865-871.
15. Ficarro, S.B., et al., *Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae.* Nature Biotechnology, 2002. **20**(3): p. 301-305.
16. Pinkse, M.W., et al., *Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns.* Analytical Chemistry, 2004. **76**(14): p. 3935-3943.
17. Larsen, M.R., et al., *Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns.* Molecular & Cellular Proteomics, 2005. **4**(7): p. 873-886.
18. Trinidad, J.C., et al., *Comprehensive identification of phosphorylation sites in postsynaptic density preparations.* Molecular & Cellular Proteomics, 2006. **5**(5): p. 914-922.
19. Fíla, J. and D. Honys, *Enrichment techniques employed in phosphoproteomics.* Amino Acids, 2012. **43**(3): p. 1025-1047.
20. Ye, J., et al., *Optimized IMAC− IMAC Protocol for Phosphopeptide Recovery from Complex Biological Samples.* Journal of Proteome Research, 2010. **9**(7): p. 3561-3573.

21. Levitt, M. and M.F. Perutz, *Aromatic rings act as hydrogen bond acceptors.* Journal of Molecular Biology, 1988. **201**(4): p. 751-754.
22. Mitchell, J.B., et al., *Amino/aromatic interactions in proteins: is the evidence stacked against hydrogen bonding?* Journal of Molecular Biology, 1994. **239**(2): p. 315-331.
23. Yaffe, M.B., *Phosphotyrosine-binding domains in signal transduction.* Nature Reviews Molecular Cell Biology, 2002. **3**(3): p. 177-186.
24. Parthasarathy, S., Y. Nishiyama, and Y. Ishii, *Sensitivity and resolution enhanced solid-state NMR for paramagnetic systems and biomolecules under very fast magic angle spinning.* Accounts of Chemical Research, 2013. **46**(9): p. 2127-2135.
25. Goldbourt, A., *Biomolecular magic-angle spinning solid-state NMR: recent methods and applications.* Current Opinion in Biotechnology, 2013.
26. Li, R., et al., *Design, synthesis, and application of a protein A mimetic.* Nature Biotechnology, 1998. **16**(2): p. 190-195.
27. Paiva, T.B., M. Tominaga, and A.C.M. Paiva, *Ionization of histamine, N-acetylhistamine, and their iodinated derivatives.* Journal of Medicinal Chemistry, 1970. **13**(4): p. 689-692.
28. Kvasnička, F. and M. Voldřich, *Determination of biogenic amines by capillary zone electrophoresis with conductometric detection.* Journal of Chromatography A, 2006. **1103**(1): p. 145-149.
29. Tsai, C.-F., et al., *Immobilized metal affinity chromatography revisited: pH/acid control toward high selectivity in phosphoproteomics.* Journal of Proteome Research, 2008. **7**(9): p. 4058-4069.
30. Ndassa, Y.M., et al., *Improved immobilized metal affinity chromatography for large-scale phosphoproteomics applications.* Journal of Proteome Research, 2006. **5**(10): p. 2789-2799.
31. Ali, A.A., et al., *Specific recognition of a multiply phosphorylated motif in the DNA repair scaffold XRCC1 by the FHA domain of human PNK.* Nucleic Acids Research, 2009. **37**(5): p. 1701-1712.
32. Nühse, T.S., et al., *Large-scale analysis of in vivo phosphorylated membrane proteins by immobilized metal ion affinity chromatography and mass spectrometry.* Molecular & Cellular Proteomics, 2003. **2**(11): p. 1234-1243.
33. Lee, J., et al., *Mitochondrial phosphoproteome revealed by an improved IMAC method and MS/MS/MS.* Molecular & Cellular Proteomics, 2007. **6**(4): p. 669-676.
34. Thingholm, T.E., et al., *SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides.* Molecular & Cellular Proteomics, 2008. **7**(4): p. 661-671.
35. Aryal, U.K., D.J. Olson, and A.R. Ross, *Optimization of immobilized gallium (III) ion affinity chromatography for selective binding and recovery of phosphopeptides from protein digests.* Journal of Biomolecular Techniques: JBT, 2008. **19**(5): p. 296.
36. Aryal, U.K. and A.R. Ross, *Enrichment and analysis of phosphopeptides under different experimental conditions using titanium dioxide affinity chromatography and mass spectrometry.* Rapid Communications in Mass Spectrometry, 2010. **24**(2): p. 219-231.
37. Imanishi, S.Y., V. Kochin, and J.E. Eriksson, *Optimization of phosphopeptide elution conditions in immobilized Fe (III) affinity chromatography.* Proteomics, 2007. **7**(2): p. 174-176.
38. Winzor, D.J., *Determination of binding constants by affinity chromatography.* Journal of Chromatography A, 2004. **1037**(1): p. 351-367.
39. Zuyi, T. and C. Taiwei, *On the applicability of the Langmuir equation to estimation of adsorption equilibrium constants on a powdered solid from aqueous solution.* Journal of Colloid and Interface Science, 2000. **231**(1): p. 8-12.
40. Winzor, D.J., *Quantitative affinity chromatography.* Journal of Biochemical and Biophysical Methods, 2001. **49**(1): p. 99-121.

41. Janin, J., *Quantifying biological specificity: the statistical mechanics of molecular recognition.* Proteins: Structure, Function, and Bioinformatics, 1996. **25**(4): p. 438-445.
42. Gohlke, H. and G. Klebe, *Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors.* Angewandte Chemie International Edition, 2002. **41**(15): p. 2644-2676.
43. Janin, J., *Principles of protein-protein recognition from structure to thermodynamics.* Biochimie, 1995. **77**(7): p. 497-505.
44. Janin, J., *Protein-protein recognition.* Progress in biophysics and molecular biology, 1995. **64**(2): p. 145-166.
45. Yan, Z., et al., *Specificity and affinity quantification of protein–protein interactions.* Bioinformatics, 2013. **29**(9): p. 1127-1133.
46. Rabe, M., D. Verdes, and S. Seeger, *Understanding protein adsorption phenomena at solid surfaces.* Advances in Colloid and Interface Science, 2011. **162**(1): p. 87-106.
47. Eriksson, A.I., et al., *Physicochemical Characterization of Phosphopeptide/Titanium Dioxide Interactions Employing the Quartz Crystal Microbalance Technique.* The Journal of Physical Chemistry B, 2013. **117**(7): p. 2019-2025.
48. Todd, R.J., R.D. Johnson, and F.H. Arnold, *Multiple-site binding interactions in metal-affinity chromatography: I. Equilibrium binding of engineered histidine-containing cytochromes c.* Journal of Chromatography A, 1994. **662**(1): p. 13-26.
49. Strandh, M., H.S. Andersson, and S. Ohlson, *Weak affinity chromatography*, in *Affinity Chromatography*2000, Springer. p. 7-23.
50. Ohlson, S. and D. Zopf (1993) *Weak-affinity chromatography*, in That T. Ngo (Ed.), Molecular Interactions in Bioseparations, Springer, p. 15-25.
51. Zhao, G., X.-Y. Dong, and Y. Sun, *Ligands for mixed-mode protein chromatography: Principles, characteristics and design.* Journal of Biotechnology, 2009. **144**(1): p. 3-11.
52. Ohlson, S., A. Lundblad, and D. Zopf, *Novel approach to affinity chromatography using "weak" monoclonal antibodies.* Analytical Biochemistry, 1988. **169**(1): p. 204-208.
53. Zopf, D. and S. Ohlson, *Weak-affinity chromatography.* Nature, 1990. **346**(6279): p. 87-88.

Chapter 3 Chromatographic evaluation and optimization of lead ligands

CHAPTER 4

Phosphopeptide enrichment using biomimetic magnetic nanostructures

Summary

Magnetic nanoparticles coated with silica and dextran (MNP-Si-Si-Dex) were successfully functionalized with lead biomimetic ligands based on triazine (ligand A8A3) and Ugi (ligand A8C2) reactions with affinity for phosphorylated peptides. The same ligands immobilized on agarose presented binding capacities 100 times lower although with higher selectivity. Four synthetic phosphopeptides consisting in consensus sequences known to bind three human phosphoprotein-binding domains were used in these studies. Ligand A8A3 immobilized on MNP-Si-Si-Dex was further used to enrich phosphopeptides from tryptic digests of α-casein, β-casein and bovine serum albumin (BSA), and the results have been compared with a Ti^{4+}-IMAC approach.

Chapter 4 Phosphopeptide enrichment using biomimetic magnetic nanostructures

4.1. Introduction

The field of nanotechnology has experienced incredible development and expansion in recent years. In particular, magnetic nanoparticles (MNPs) present the attractive feature of manipulation by an external magnetic field, finding applications in countless fields of science, such as medical imaging, drug delivery, biosensing, bioseparation, among others [1-5]. This class of nanoparticles includes metallic (e.g. Fe, Co, and Ni nanoparticles), bimetallic (e.g. FePt, FeCo), and iron oxide MNPs [1, 6]. The low toxicity, ease of synthesis, and biodegradability of the latter have boosted their widespread utilization, and in fact a few formulations have already been approved for clinical usage [1, 6]. In the field of magnetic separations, iron oxide MNPs have been used since the 1960´s [4]. By selecting the appropriate method of synthesis and coating agents, it is possible to tailor their size, shape, and surface chemistry [1, 6]. As a consequence of their small size, MNPs have a high surface-to-volume ratio and tend to agglomerate to decrease their surface energy [3, 7]. Due to their highly reactive surfaces, MNPs are generally coated with organic and/or inorganic materials to protect them against oxidation and acid/base erosion [2, 7]. The choice of the coating agent is extremely important as it allows tuning the surface charge, hydrophobic/hydrophilic character, and chemical functionality [6]. Polymeric coatings have been widely used to prevent particle aggregation and non-specific binding, but the physical-chemical characteristics of the polymer (e.g. molecular weight, hydrophobicity, chemical structure, biodegradability) must be considered, as they may affect particles' performance [1, 6, 8]. In the last decade, MNPs have been extensively used for the enrichment of phosphorylated peptides. Magnetite (Fe_3O_4) was successfully used for the selective capture of phosphopeptides from tryptic digests of bovine β-casein, cytochrome c, bovine serum albumin (BSA), and horse heart myoglobin, but no enrichment was observed when maghemite (γ-Fe_2O_3) was used [9]. As Fe_3O_4 is easily converted to γ-Fe_2O_3 in the presence of oxygen [3], one option is to modify the surface of the particles, minimizing oxidation and introducing functional groups with high affinity for phosphopeptides. Coating agents range from metal oxides (e.g. Al_2O_3, ZrO_2, TiO_2, Ga_2O_3, SnO_2, Ta_2O_5, Nb_2O_5, ZnO) [10], polymers (e.g. polyethyleneimine) [11], hydroxyapatite [12], zirconium arsenate [13], yttrium phosphate [14], Ti^{4+}-adenosine tri-phosphate [15], lanthanum silicate [16], rare earth vanadates ($REVO_4$; RE = Sm, Dy, Ho) [17], bentonite [18], ammonium fluoride and lutetium fluoride [19], guanidinium [20], alkoxide-bridged dinuclear zinc(II) complex of 1,3-bis[bis(pyridine-2-ylmethyl)amino]propan-2-olate (Zn^{2+}-Phos-tag) [21], among many others. The binding mode of the majority of these surface-modified MNPs is still by metal chelation. In this work, a different approach is developed by using small biomimetic ligands immobilized on MNPs coated with silica and dextran.

Silica is used in order to protect MNPs surface and prevent iron leaching pH at acidic pH, while dextran is a biocompatible polymer composed of glucose units and is used to minimize non-specific binding.

4.2. Materials and Methods

4.2.1. Chemicals

All reagents were of the highest purity available and the solvents were pro-analysis.

1-Pyrenemethylamine hydrochloride (AP), 2,5-dihydroxybenzoic acid (2,5-DHB), 3-aminopropyltriethoxy silane (APTES), ammonium bicarbonate (NH_4HCO_3), ammonium hydroxide solution (NH_4OH), cyanuric chloride, dextran from *Leuconostoc mesenteroides* (average mol wt 150,000), DL-dithiothreitol (DTT), epichlorohydrin, formic acid, glutaraldehyde solution (50 wt. % in H_2O), histamine (A8), iodoacetamide (IAA), iron(II) chloride tetrahydrate, iron (III) chloride hexahydrate, isopropyl isocyanide, ninhydrin, phenethylamine (A3), phenol, phosphoric acid (H_3PO_4), potassium cyanide, pyridine, sodium metasilicate pentahydrate, succinamic acid (C2), tetraethyl orthosilicate (TEOS), trifluoroacetic acid (TFA), were acquired from Sigma-Aldrich. 2-Propanol and methanol (MeOH) were obtained from Roth, glycine from Acros Organics, and acetic acid glacial from Pronalab. Acetone and dimethylformamide (DMF) were purchased from VWR. acetonitrile (ACN), ethanol (EtOH) absolute PA, hydrochloric acid (HCl) 37%, sodium chloride (NaCl) and sodium hydroxide (NaOH) were acquired from Panreac. Urea was acquired from Merck. Nitrogen gas was provided by Air Liquide. Ser-Gln-Val-Phe-Pro-Trp (SW6), pSer-Gln-Val-Phe-Pro-Trp (SW6-P), Thr-Gln-Val-Asp-Ala-Trp (TW6), pThr-Gln-Val-Asp-Ala-Trp (TW6-P), Tyr-Glu-Glu-Ile-Pro-Trp (YW6), pTyr-Glu-Glu-Ile-Pro-Trp (YW6-P), Tyr-Ala-Gly-Ser-Thr-Asp-Glu-Asn-Thr-Asp-Ser-Glu-Trp (YW13), and Tyr-Ala-Gly-pSer-pThr-Asp-Glu-Asn-pThr-Asp-Ser-Glu-Trp (YW13-P) peptides were > 98% pure and were purchased from Genecust and Caslo.

4.2.2. Biochemicals

Trypsin Gold, Mass Spectrometry Grade was purchased from Promega. α-Casein from bovine milk (Cat. No. C6780), β-Casein from bovine milk (Cat. No. C6905), and Bovine Serum Albumin (BSA) (Cat. No. A2153) were acquired from Sigma-Aldrich.

4.2.3. Materials

Half-area UV-Star® 96-well microplates, black immunograde 96-well microplates, and 96-well transparent microplates were obtained from Greiner Bio-One, Brand, and Sarstedt, respectively. Sep-Pak C18 cartridges were purchased from Waters. GELoader tips were purchased from Eppendorf.

4.2.4. Instrumentation

Ligand synthesis was carried out in an IKA KS 4000 ic control shaker. Samples were incubated with ligand-functionalized MNP-Si-Si-Dex in a VorTemp^{TM} 56 shaking incubator from Labnet. Fluorescence readings in 96-well microplates were conducted in a Tecan F200 Microplate Reader using a λexc = 280 (20) nm – λem = 340 (35) nm filter, and 560 (10) nm filter, and a 280 (5) nm filter (Tecan). Dynamic light scattering (DLS) analysis was conducted on a Zetasizer Nano ZS from Malvern. All samples were prepared in Milli-Q water to a final concentration of 0.05 mg/mL.
Prior to Matrix Assisted Laser Desorption/Ionization Time-Of-Flight Mass Spectrometry (MALDI-TOF MS) analysis, the samples were resuspended in 10 µL 2,5-DHB (25 mg/mL prepared in 70% (v/v) MeOH/ 1% (w/v) H_3PO_4), 0.5 µL applied on a MALDI sample plate and allowed to dry. MS analysis was performed on a Waters MALDI micro MX system (Waters, Milford, MA, USA) equipped with a nitrogen UV laser (337 nm). Measurements were done in the positive reflection mode with a reflectron voltage of 5200 V, delayed extraction time of 700 ns, pulse voltage of 2100 V, flight tube voltage of 12000 V, and a detector voltage between 2300-2600 V. An average of 750 laser shots were acquired per spectrum. The spectra were processed (baseline correction, smoothing and deisotoping) with the Masslynx software (Waters) and the database searches were performed using the Mascot search engine.

4.2.5. Methods

4.2.5.1. Synthesis of iron oxide magnetic nanoparticles (MNPs)

A 150 mL aqueous solution of 2 M NH_4OH was purged with nitrogen gas for 30 min in a closed reactor equipped with a mechanical stirrer (500 rpm). A 50 mL aqueous solution containing 5.4 g of $FeCl_3.6H_2O$ and 2 g of $FeCl_2.4H_2O$ was added dropwise to the reactor and the reaction proceeded for

2 h with agitation (500 rpm) and under inert atmosphere. Finally, the particles were thoroughly washed with distilled deionised water (dd H_2O) using magnetic decantation.

4.2.5.2. Coating of MNPs with two silica layers and dextran (MNP-Si-Si-Dex)

An aqueous solution of MNPs (10 mg/mL, 190 mL) was placed in an ultrasonic bath for 15 min at 37 kHz. Afterwards, 2.28 g of sodium silicate in 80 mL of 50% (v/v) EtOH/dd H_2O was added to the MNPs and the solution was incubated in a water bath at 40°C with mechanical stirring (500 rpm) for 2 h. In the end, particles coated with a silica layer (MNP-Si) were washed by magnetic decantation with 50% (v/v) EtOH/dd H_2O (2x 270 mL) and dd H_2O (3x 270 mL) to remove the excess of sodium silicate. MNP-Si were then resuspended in 190 mL of dd H_2O and an aliquot of 5 mL was taken for further characterization. A solution of 1.85 g MNP-Si in 80 mL of 80% (v/v) EtOH/dd H_2O was ultrasonicated for 10 min at 37 kHz, followed by addition of 3 mL 5 M NH_4OH. TEOS (1.5 mL) was then added slowly to the previous solution under ultrasonication for 10 min at 37 kHz. The solution was incubated in a water bath for 2h at 40°C with mechanical stirring (500 rpm). The resulting particles with two layers of silica (MNP-Si-Si) were washed with 80% (v/v) EtOH/dd H_2O (2x 185 mL) and dd H_2O (3x 185 mL) by centrifugation (13500 *g*, 20 min), resuspended in 185 mL dd H_2O, and an aliquot of 5 mL was taken for characterization. An aqueous solution of MNP-Si-Si (14.4 mg/mL, 125 mL) was ultrasonicated for 5 min at 37 kHz, followed by the addition of 125 mL of an aqueous solution of dextran (14.4 mg/mL). The mixture was then incubated in a water bath at 60°C for 2 h with mechanical stirring (500 rpm). Finally the particles coated with silica and dextran (MNP-Si-Si-Dex) were washed with dd H_2O (5x 250 mL) by centrifugation (13500 *g*, 20 min) and resuspended in dd H_2O.

4.2.5.3. Amination of MNP-Si-Si-Dex

A solution of MNP-Si-Si-Dex (8mg/mL, 179 mL) in 50% (v/v) EtOH/dd H_2O was ultrasonicated for 5 min at 37 kHz in an ultrasonic bath, followed by the addition of 1 mL acetic acid glacial and 20 mL APTES. The mixture was incubated in an orbital shaker (200 rpm) for 1h at 70°C. The particles were washed with 50% (v/v) EtOH/dd H_2O (3x 200 mL) and dd H_2O (3x 200 mL) by magnetic decantation. Kaiser test was performed according to §2.2.4.4, but 1 mL of aminated particles (MNP-Si-Si-Dex-NH_2) was used instead of the moist aminated agarose. Gly solutions (0-0.005 M) were used as standards. After the heating step, both samples and standards were diluted 1:20 and their absorbance was determined in a microplate reader using a 560 (10) nm filter. Samples were centrifuged to remove

the particles from solution before absorbance measurement. The amount of amine groups was estimated to be 529 ± 41 μmol NH_2/ g MNP-Si-Si-Dex-NH_2.

4.2.5.4. Functionalization of MNP-Si-Si-Dex-NH_2 with aldehyde groups

MNP-Si-Si-Dex-NH_2 (0.34 g) were suspended in 100 mL of 5% (v/v) glutaraldehyde/dd H_2O, 6 mL of NaOH 1 M were added, and the solution was ultrasonicated for 5 min at 37 kHz, followed by incubation in an orbital shaker (250 rpm) for 1 h at 30°C. In the end, aldehyde functionalized particles (MNP-Si-Si-Dex-CHO) were washed with dd H_2O (6x 100 mL) by magnetic decantation.

4.2.5.5. Synthesis of triazine ligands (APAP and A8A3), and Ugi ligands (APC2 and A8C2) on MNP-Si-Si-Dex

Triazine ligands were synthesized according to §2.2.4.10 and §3.2.5.3, and Ugi ligands were synthesized according to §2.2.4.9 and §3.2.5.2. The molar equivalents of each reactant were determined relative to the amount of amine groups on MNP-Si-Si-Dex-NH_2, instead of epoxy groups on epoxy-activated agarose.

4.2.5.6. Screening of biomimetic ligands against phosphorylated peptides

Microcentrifuge tubes containing 10 mg of MNP-Si-Si-Dex, MNP-Si-Si-Dex functionalized with ligand A8A3 (MNP-Si-Si-Dex-A8A3) or MNP-Si-Si-Dex functionalized with ligand A8C2 (MNP-Si-Si-Dex-A8C2) were placed in a 24-tube magnetic separation rack. The particles were washed with 0.1 M NaOH in 30% (v/v) isopropanol followed by dd H_2O (3 cycles of washes, 1 mL/ tube), and afterwards with 50 mM Gly pH 3 (8x 1 mL/ tube). The process consists of adding the solvent/ buffer to the particles in the tubes, vortex the solution, place the tubes in the rack, wait until the particles are attracted to the magnet and the solution is clear, and remove the supernatant. A mixture of non-phosphorylated peptides (SW6, TW6, and YW6; 0.5 mg/mL total peptide) and a mixture of phosphorylated peptides (SW6-P, TW6-P, and YW6-P; 0.5 mg/mL total peptide) in 50 mM Gly pH 3 were prepared. YW13 and YW13-P were dissolved either in 50 mM Gly pH3 50 mM NaCl (for screening of MNP-Si-Si-Dex-A8A3) or 50 mM Gly pH3 100 mM NaCl (for screening of MNP-Si-Si-Dex-A8C2), for a final concentration of 0.5 mg/mL. Each particle-containing tube was incubated with 1 mL of each peptide solution for 1h at room temperature in a shaker incubator (1200 rpm). After that period, the tubes were placed back in

the rack and the supernatant removed. The particles were then washed with their correspondent binding buffer (6x 1mL/ tube). Both the flow-through and washes were pipetted into the wells of 96-well black microplates, and the fluorescence intensity was read on a microplate reader using a λexc = 280 (20) nm – λem = 340 (35) nm filter.

4.2.5.7. Preparation of protein digests

α-Casein and β-casein were solubilized in 50 mM NH_4HCO_3 pH 8.2 to a final concentration of 1 mg/mL), and incubated with trypsin (1:100 enzyme-to-protein ratio (w/w)) for 12h at 37°C with gentle agitation. BSA was dissolved in 50 mM NH_4HCO_3 pH 8.2 with 8 M urea (1 mg/mL). DTT was added to a final concentration of 5 mM and the mixture was incubated for 1h at 37°C with mild agitation. IAA was added to a final concentration of 10 mM and the BSA solution was incubated in the dark for 30 min at room temperature. DTT was added to a final concentration of 5 mM to quench excess IAA and prevent overalkylation and the BSA solution was incubated in the dark for another 30 min at room temperature. The solution was diluted 5 times with 50 mM NH_4HCO_3 pH 8.2 to reduce the urea concentration, and then incubated with trypsin (1:100 enzyme-to-protein ratio (w/w)) for 12h at 37°C with gentle agitation. 1% TFA (v/v) was then added to stop enzymatic activity. Each protein digest was desalted using Sep-Pak C18 cartridges. Sep-Pak columns were conditioned with ACN (2x 1 mL), and then equilibrated with 0.6% (v/v) acetic acid (2x 1 mL). The acidified tryptic digests (1 mL) were loaded onto the columns, the columns were washed with 0.6% (v/v) acetic acid (3x 1 mL), and the peptides were eluted with 80% ACN/ 0.6% acetic acid (v/v) (2x 500 µL). The elution fractions were pooled together, aliquoted, dried down in the SpeedVac and stored at -80°C. The desalting step was performed in an extraction manifold connected to a vacuum pump.

4.2.5.8. Phosphopeptide enrichment using Ti^{4+}-IMAC and semi-complex protein digests

A 534 µm C8 Empore disk was introduced into a GELoader spin tip using a 365 µm (outer diameter) fused silica capillary, and the tip was end-cut 0.2 µm below the C8 plug. The column was washed with 20 µL MeOH. Ti^{4+}-IMAC slurry (10 mg/mL in 30% ACN/ 0.1% TFA (v/v)) was pipetted into the GELoader tip, and the tip was inserted into a microcentrifuge tube equipped with an adaptor. The tube was centrifuged at 100 *g* in order to properly pack the Ti^{4+}-IMAC microcolumn. Ti^{4+}-IMAC material had been previously prepared according to [22, 23]. Ti^{4+}-IMAC GELoader spin tip was

equilibrated with 50 µL of loading buffer (80% ACN/6% TFA (v/v)) by centrifugation at 100 *g* for 5 min. A semi-complex protein sample consisting of tryptic digests of α-casein, β-casein and BSA (1:1:1) was prepared in loading buffer, loaded onto the Ti^{4+}-IMAC spin tip (0.12 pmol/µL; 50 µL), and centrifuged at 50 *g* for 30 min. The Ti^{4+}-IMAC spin tip was then washed with 50 µL of washing buffer 1 (50% ACN/0.5% TFA, 200 mM NaCl) by centrifugation at 170 *g* for 17 min, followed by 50 µL of washing buffer 2 (50% ACN/0.1% TFA) by centrifugation at 170 *g* for another 17 min. Phosphopeptides were eluted into a microcentrifuge tube containing 35 µL of 10% (v/v) formic acid, by adding 20 µL of 10% (v/v) NH_4OH into the Ti^{4+}-IMAC spin tip, followed by centrifugation at 100 *g* for 20 min. Samples from flow-through and washes were dried down in a SpeedVac and stored at -80°C for further MS analysis.

4.2.5.9. Phosphopeptide enrichment using biomimetic ligands A8A3 and A8C2 and semi-complex protein digests

Microcentrifuge tubes containing 5 mg of ligand-functionalized MNP-Si-Si-Dex (either MNP-Si-Si-Dex-A8A3 or MNP-Si-Si-Dex-A8C2) were conditioned with loading buffer (2x 500 µL) in a shaker incubator (1200 rpm) for 10 min at room temperature. Then, the tubes were placed in a 24-tube magnetic separation rack, and supernatant was removed. Semi-complex protein digest samples of α-casein, β-casein and BSA (1:1:1) (0.024 pmol/µL; 250 µL) were loaded into the tubes and incubated for 1h at room temperature with agitation (1200 rpm). The supernatant was removed and the particles were washed first with washing buffer 1, and then with washing buffer 2. The washing steps consisted in incubating the particles with 250 µL of each buffer in a shaker incubator (1200 rpm) for 10 min at room temperature. The elution step consisted in incubating the particles with 10% (v/v) NH_4OH for 10 min at room temperature with agitation. The elution fraction was immediately dried down in a SpeedVac and stored at -80ºC for further MS analysis.

4.3. Results and Discussion

4.3.1. Synthesis of triazine and Ugi based ligands on MNP-Si-Si-Dex

The feasibility of triazine and Ugi reactions on magnetic nanoparticles coated with two layers of silica and dextran was assessed by using pyrene as a reporter molecule. 1-Pyrenemethylamine (AP) was used as the amine component in both triazine and Ugi reactions (Figure 4.1).

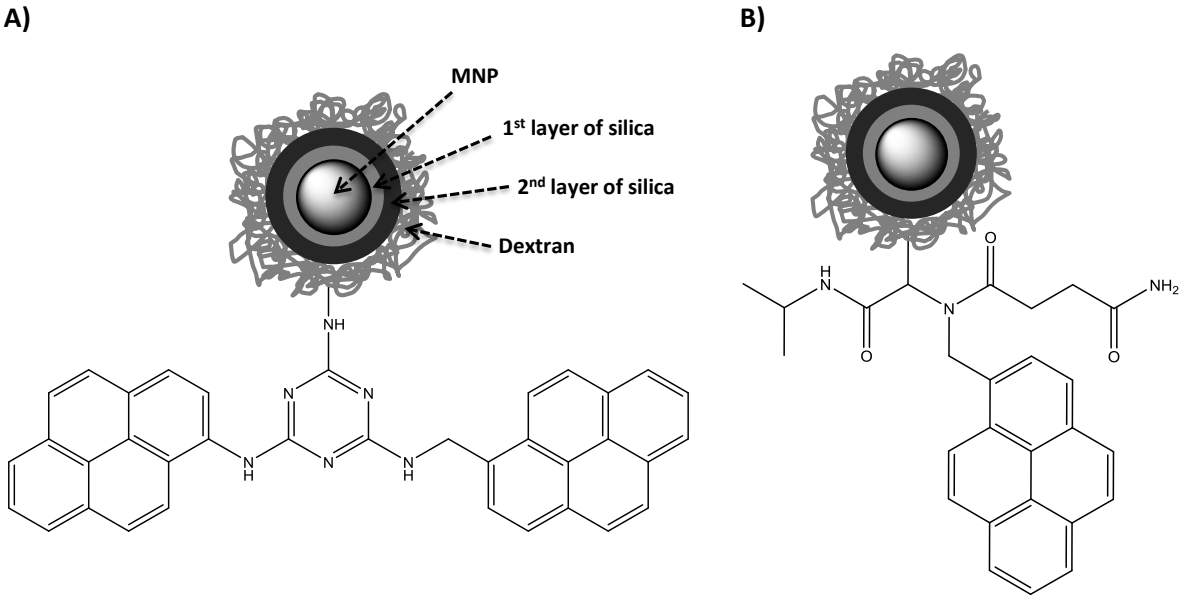


Figure 4.1 - Schematic representation of ligands A) APAP (triazine-based) and B) APC2 (Ugi-based) immobilized onto MNP-Si-Si-Dex.

MNP-Si-Si-Dex were aminated using APTES (MNP-Si-Si-Dex-NH_2) and directly used for the synthesis of triazine ligand APAP. Aldehyde groups were further introduced at the surface of the particles by reacting the aminated particles with glutaraldehyde, and MNP-Si-Si-Dex-CHO were used as the aldehyde component in the Ugi reaction. Fluorescence microscopy results revealed that the fluorescence intensity for the same exposition time was lower for MNP-Si-Si-Dex modified with triazine ligand APAP, when compared to MNP-Si-Si-Dex functionalized with Ugi ligand APC2 (Figure 4.2). Fluorescence images were analyzed in ImageJ 1.42q software by measuring maximum gray values in selected particle-containing areas. For an 8 bits per pixel image, the intensity of each pixel is represented by a number between 0 and 255, where 0 is black, 255 is white, and the values in between correspond to different shades of gray. Although ligand APAP has two pyrene molecules in its structure, APC2 presents higher fluorescence intensity, which is translated into a gray value two

times higher (Figure 4.3). This might be due to: (i) a higher density of APC2 ligand at the surface of MNP-Si-Si-Dex; or (ii) crosslinking effect of glutaraldehyde during APC2 ligand synthesis causing particle aggregation [24]. Glutaraldehyde is a well-known protein crosslinking agent, readily reacting with nucleophiles, such as amines, thiols, phenols, etc. It might exist under monomeric, polymeric, linear and cyclic forms, and its exact mechanism of action is subject of scientific debate [25]. The crosslinking effect may also arise from the presence of unreacted primary amines at the surface of the solid support, which can react with aldehyde groups also at the surface, instead of the amine component in solution. An alternative methodology consists in coating the particles with oxidized dextran, which would directly introduce aldehyde functional groups at the surface, and would avoid the crosslinking effect of glutaraldehyde [26, 27].

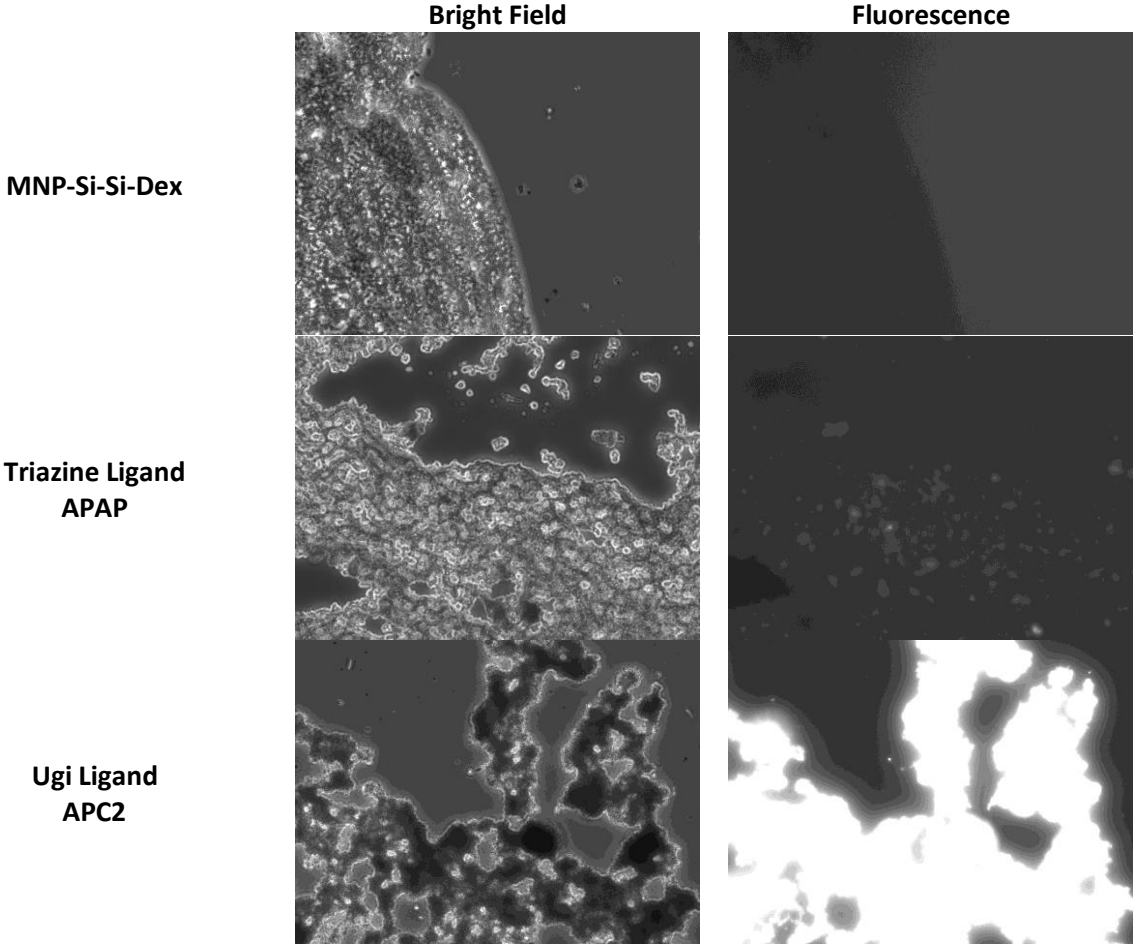


Figure 4.2 – Fluorescence microscopy results of triazine and Ugi scaffolded MNP-Si-Si-Dex functionalized with a fluorophore – pyrene. MNP-Si-Si-Dex was used as negative control. The images were acquired using an Olympus BX51 microscope (400X magnification) with a U-MWB filter (λexc 460–490 nm; λexc 515–700 nm) and an Olympus U-RFL-T lamp. All images were acquired with an exposition time of 5 s. The monitoring software was Cell-View System.

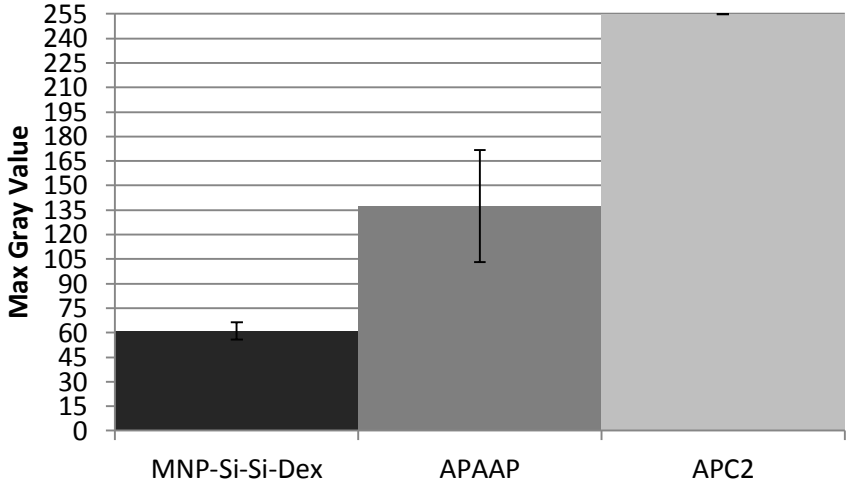


Figure 4.3 – Maximum gray values for MNP-Si-Si-Dex functionalized with ligands APAP and APC2 measured using ImageJ 1.42q software. (N=10)

Fluorescence microscopy results indicate that MNP-Si-Si-Dex is suitable for the solid-phase synthesis of both triazine and Ugi ligands. Therefore, triazine (A8A3) and Ugi (A8C2) phosphopeptide-binding ligands identified in the previous chapter were synthesized on MNP-Si-Si-Dex (Figure 4.4).

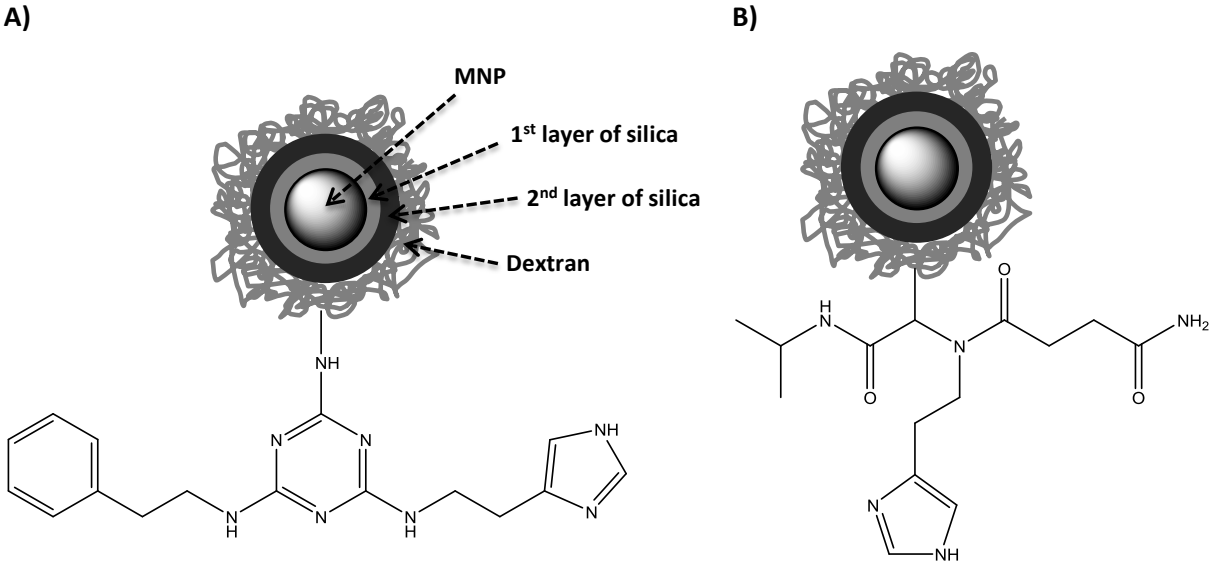


Figure 4.4 – Schematic representation of ligands A) A8A3 and B) A8C2 immobilized onto MNP-Si-Si-Dex.

The magnetic support was analysed by DLS after each reaction step (Figure 4.5). DLS allows the determination of the size of molecules and particles (typically in the sub-micrometre region) by measuring their Brownian motion in solution. In simple terms, particles in suspension are in constant movement, and they scatter the light at different intensities upon laser illumination. This allows the

determination of the velocity of the Brownian motion, and consequently the size of the particles using the Stokes-Einstein equation [5] (smaller particles move faster than larger particles) [29, 30].

The size of each individual non-hydrated particle is accurately determined by transmission electron microscopy (TEM), while DLS allows the determination of hydrodynamic diameters (HDs). HDs have been measured for aqueous samples with a final concentration of 0.05 mg/ml, and refer to the size of solvated particles. Bare MNPs synthesized by the co-precipitation method presented herein have core diameters of approximately 10 nm [31], but they rapidly aggregate in solution and even during sample analysis by DLS, presenting HDs of approximately 2.3 µm (Figure 4.5A). Coating the particles with the first layer of silica using sodium silicate (MNP-Si) is not sufficient to disperse them in solution, as the HDs continue to present values in the micrometre region. However, when the particles are coated with TEOS (MNP-Si-Si), the second layer of silica makes them highly hydrophilic and easily dispersed in aqueous solution, which is translated into HDs of approximately 177 nm. Taylor and co-workers have already reported that one layer of silica is insufficient to completely shield MNPs surface. In addition, TEOS provided higher silica content when compared to sodium silicate [32]. Modification with silica not only boosts particles dispersion in solution, but also improves their resistance against iron leaching at acidic environments [33]. Dextran modification (MNP-Si-Si-Dex) did not have any significant effect on the HDs of the silica-coated particles; therefore, it is uncertain if the amount of dextran used was enough to completely coat the surface. Most likely, dextran and silica formed an interconnected porous network and, being dextran a neutral polysaccharide, the negative charges of silica prevail in inter-particle interactions. By coating iron oxide particles with dextran strands crosslinked with silica, Grüttner and Teller were able to increase their robustness and mechanical properties [34]. Particles amination using APTES (MNP-Si-Si-Dex-NH_2) increases their HD, probably due to non-covalent interactions between amine and hydroxyl groups of neighbour particles. This effect has already been observed for dextran coated MNPs [35]. Functionalization of MNP-Si-Si-Dex-NH_2 with triazine ligand A8A3 is performed by two subsequent nucleophilic substitutions on the triazine ring. After the first substitution with histamine (Ligand A8), there is a decrease in the HD of the particles, probably due to mixed-mode interactions provided by the triazine ring and the imidazole group of histamine, which provide both hydrophobicity through their aromatic groups and hydrophilicity through their numerous nitrogen atoms. After the second nucleophilic substitution with phenethylamine (Ligand A8A3), the higher degree of hydrophobicity leads to larger aggregates in aqueous solution. Particles coated with

[5] $HD = \frac{k_B T}{3 \pi \eta D_t}$, where HD is the hydrodynamic diameter, k_B is Boltzmann's constant, T is the thermodynamic temperature, η is the dynamic viscosity, and D_t is the translational diffusion coefficient [28].

glutaraldehyde (MNP-Si-Si-Dex-CHO) present the highest HD values. These particles are still agglomerated after the Ugi reaction (Ligand A8C2), corroborating the crosslinking effect of glutaraldehyde discussed previously.

Zeta potential values were also determined for all samples and are presented in Figure 4.5B. A positively or negatively charged particle dispersed in solution will attract oppositely charged ions to its surface. Ions close to the surface will be strongly bound, forming the so called Stern layer, while ions more distant to the surface are loosely bound and constitute the diffuse layer. Within this layer there is a notional boundary coined the slipping plane and any ions within it move along with the particle, while ions outside this boundary will not move with the particle. The potential between the particle surface and the dispersing liquid at the slipping plane is called zeta potential (Figure 4.6) [36]. Zeta potential values give an indication about the charge of the particles and their tendency to flocculate. Particles are stable in solution when they present zeta potential values above 30 mV or below -30 mV.

Bare MNPs presented zeta potential values of approximately 5.6 mV, which means they have tendency to flocculate in solution (Figure 4.5B). The modification with the first layer of silica did not have any significant effect on the zeta potential, which is consistent with the large HDs observed for both MNP and MNP-Si. The second layer of silica (MNP-Si-Si) drastically decreased particles' zeta potential to -30.7 mV, which is in accordance with the small HDs, as particles are stable in solution. This effect has been observed by Sun and co-workers, who reported that increasing the amount of TEOS, i.e. increasing the thickness of the silica layer, helps sheltering the magnetic dipole attraction and dispersing particles in solution. In addition, the silica layer is negatively charged at neutral pH, increasing the coulomb repulsion between the particles [37]. The negative value of MNP-Si-Si-Dex zeta potential corroborates the existence of silica at the surface, either due to insufficient amount of dextran at the surface or to the existence of a polymer-silica network shell. MNP-Si-Si-Dex-NH_2 presents a zeta potential value of -7.2 mV, a much less negative value than MNP-Si-Si-Dex, due to the existence of amine groups at the surface. This was further confirmed by the Kaiser test, a colorimetric test, which gave a characteristic purple/blue colour owing to the presence of primary amines. Although it is essentially a qualitative test, the amount of amine groups was estimated to be 529±41 μmol NH_2/ g support. Ligand A8 presents a positive zeta potential value (15.6 mV) due to the positive charges on the triazine ring and imidazole group of histamine. However, the zeta potential after the second nucleophilic substitution decreases to -4.7 mV, a value close to zero, which indicates that the particles tend to form bigger aggregates. This is in accordance with the larger HDs observed for ligand A8A3 when compared to ligand A8. MNP-Si-Si-Dex-CHO present zeta potential values

more negative than MNP-Si-Si-Dex-NH_2, which means the reaction with glutaraldehyde was successful, and at least some of the amine groups reacted to give terminal aldehyde functional groups. Ligand A8C2 presents positive zeta potential due to the presence of histamine. However, it should be kept in mind that the interpretation of zeta potential values is complex, especially when referring to polymer- or ligand-coated particles. A multiplicity of factors, such as ionic strength, coating thickness, surface curvature, and polymer permeability, must be considered [38].

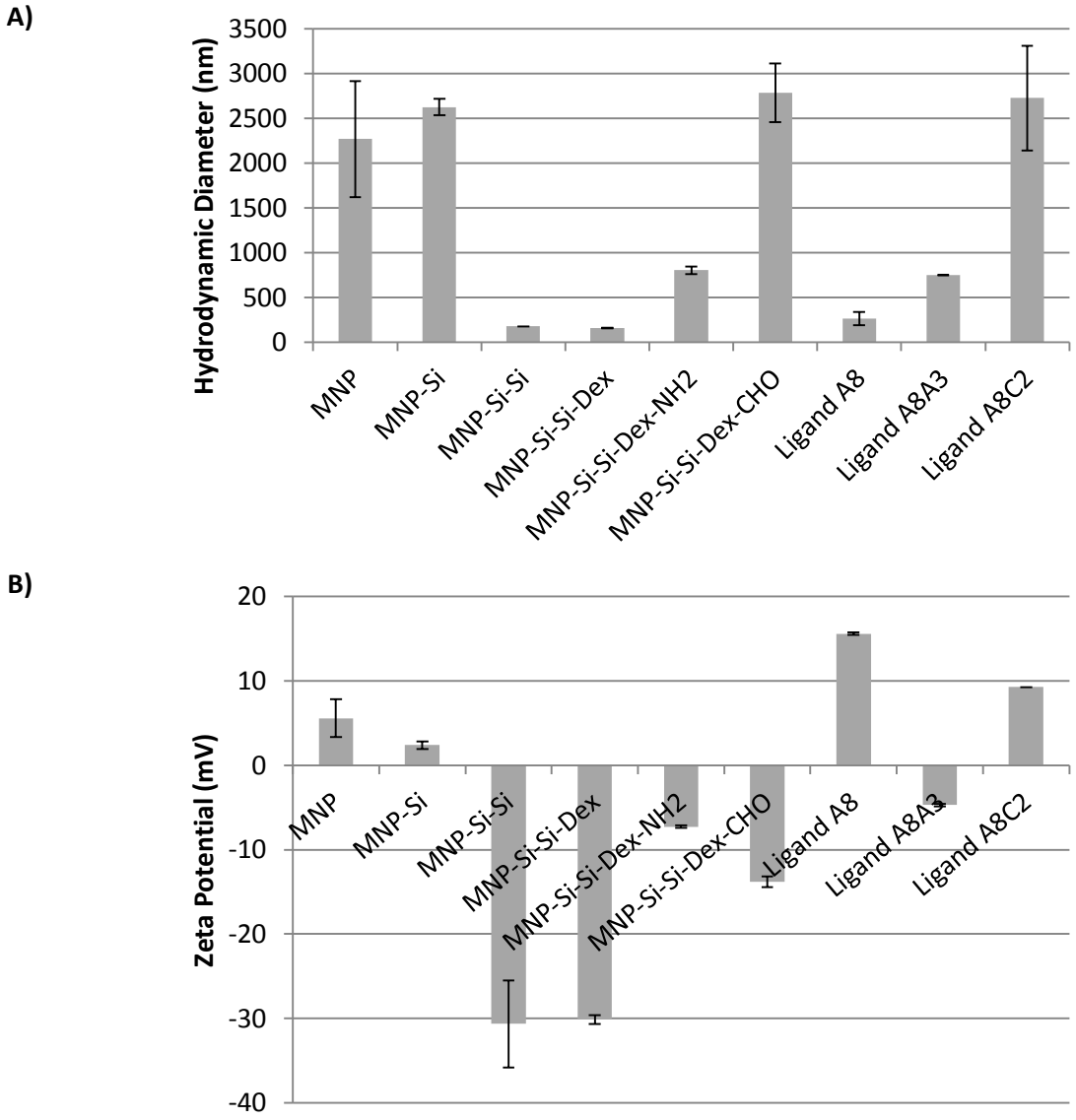


Figure 4.5 – DLS results of triazine and Ugi-modified particles. A) Hydrodynamic diameters (nm). B) Zeta potential values (mV). All samples have a concentration of 0.05 mg/ml in Milli-Q water and the analysis was conducted on a Zetasizer Nano ZS from Malvern.

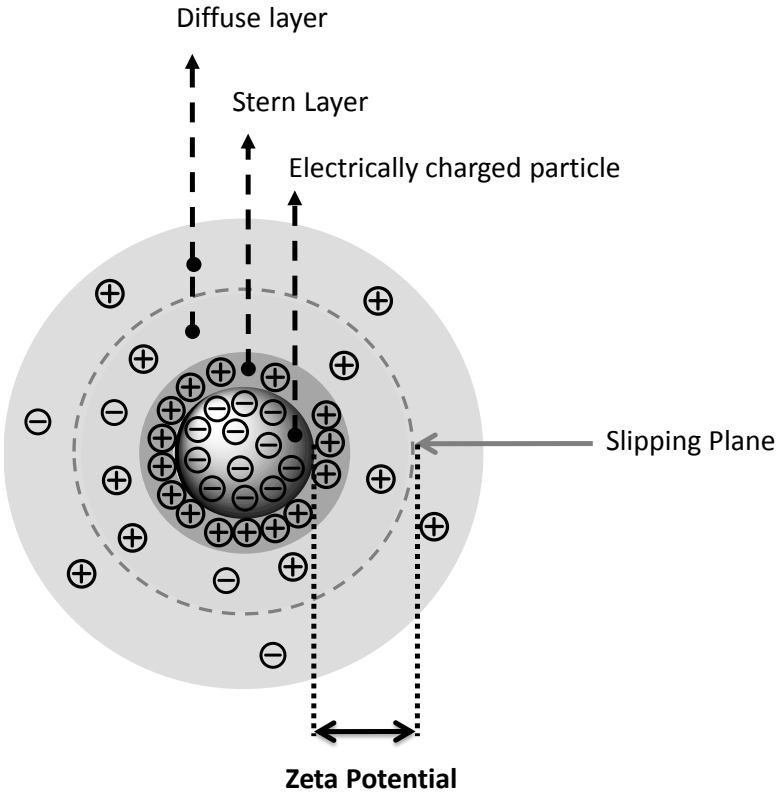


Figure 4.6 – Schematic representation of zeta potential. When a negatively charged particle is diffusing in solution, a high concentration of counterions will be attracted and strongly attached to the surface (Stern layer), while ions located farther from the surface will be more loosely bound (diffuse layer). The diffuse layer comprises an inner region, consisting of ions travelling along with the particle upon its movement, and an outer region, comprising ions which will not move when the particle moves. The boundary between these regions is called the slipping plane, and the potential measured at the slipping plane is called zeta potential.

4.3.2. Comparison between phosphopeptide-binding affinity ligands immobilized in agarose and MNP-Si-Si-Dex

The binding capacity and phosphopeptide enrichment efficiency of ligands A8A3 and A8C2 immobilized on MNP-Si-Si-Dex have been determined by screening the ligands against a mixture of mono-phosphorylated peptides (SW6-P, TW6-P, and YW6-P) and the multi-phosphorylated peptide YW13-P. The results were compared with the same ligands immobilized on agarose, which have been reported on the previous chapters, and are presented in Table 4.1. Generally speaking, the binding capacities of the ligands immobilized on MNP-Si-Si-Dex increase 100-fold on average when compared to the same ligands immobilized on agarose, but at the cost of lower phosphopeptide enrichment efficiencies, which are translated into lower phosphopeptide mass fractions. The higher binding capacity of ligand-functionalized MNP-Si-Si-Dex is probably due to the higher amount of functional

groups available at the surface of the particles, since the support itself (non-functionalized MNP-Si-Si-Dex) presents a maximum of 20% non-specific adsorption. Nonetheless, the effect of non-specific interactions should not be discarded, as they may contribute for the lower selectivity of the ligand-functionalized support towards the phosphorylated targets.

The same ligand (either A8A3 or A8C2) present similar binding capacities values for both SW6-P+TW6-P+YW6-P mixture and YW13-P, but the enrichment efficiency is always superior for the multi-phosphorylated peptide. This is valid for ligands immobilized on both solid supports (Table 4.1).

Table 4.1 – Screening results of ligands A8A3 and A8C2 immobilized either on agarose or MNP-Si-Si-Dex against a mixture of mono-phosphorylated peptides (SW6-P, TW6-P, and YW6-P) and a multi-phosphorylated peptide (YW13-P). Screening was performed by loading 0.25 ml of 1 mg/ml phosphorylated peptides in 50 mM Gly pH3. The supports were washed with the same buffer until no peptide was detected in the recovered fractions. The samples were analysed by measuring the florescence intensity in a microplate reader equipped with a λexc = 280 (20) nm – λem = 340 (35) nm filter. The results were normalized per gram of support. BC is the binding capacity in mg target peptide per g support; PMF is the phosphopeptide mass fraction in %.

		Solid Support							
		Agarose				MNP-Si-Si-Dex			
		A8A3		A8C2		A8A3		A8C2	
		BC (mg/g)	PMF (%)	BC (mg/g)	PMF (%)	BC (mg/g)	PMF (%)	BC (mg/g)	PMF (%)
Target Peptides	SW6-P+TW6-P+YW6-P	0.35 ±0.01	91.83 ±0.02	0.32 ±0.01	87.60 ±0.61	32.61 ±0.44	64.45 ±13.74	43.92 ±2.23	62.23 ±1.30
	YW13-P	0.31 ±0.01	98.15 ±2.61	0.39 ±0.01	99.64 ±0.51	37.09 ±1.63	81.74 ±3.28	42.80 ±0.10	72.09 ±1.18

4.3.3. Comparative evaluation of phosphopeptide-binding biomimetic ligands with a novel Ti^{4+}-IMAC approach for phosphopeptide enrichment

The specificity of ligands A8A3 and A8C2 towards phosphopeptides was further tested by screening the ligands against a semi-complex mixture composed of tryptic digests of two standard phosphorylated proteins (α-casein and β-casein) and one non-phosphorylated protein (BSA). The eluted peptides have been characterized by MALDI-TOF MS, and the enrichment efficiency of the synthetic ligands has been compared with a novel Ti^{4+}-IMAC adsorbent developed by Zhou and co-workers [22, 39]. The same loading, washing and elution buffers have been used in both experiments.

The Ti^{4+}-IMAC adsorbent consists of monodisperse polystyrene microspheres functionalized with a flexible poly(GMA-co-TMPTMA) linker (GMA is glycidyl methacrylate and TMPTMA is trimethylol-

propane trimethacrylate) coupled to phosphonate groups, which in turn coordinate the Ti^{4+} (Figure 4.7). These microspheres are stable under strong acid and alkaline conditions, have large surfaces areas, a hydrophilic surface to minimize non-specific adsorption, and have uniform monodisperse size distribution, which allows homogeneous column packing, uniform flow profile, and low column pressure. The poly(GMA-co-TMPTMA) spacer provides an enhanced hydrophilic character to the support, besides diminishing steric hindrance effects. Phosphonate groups are used as chelating ligands, instead of the commonly used iminodiacetic acid (IDA) and nitriloacetic acid (NTA), and are reported to provide an advantageous structural orientation [22, 39].

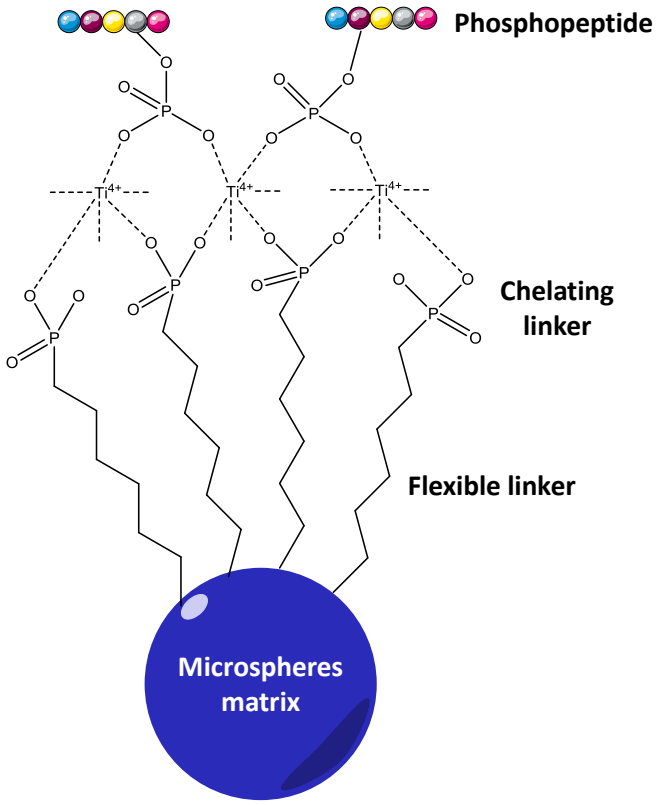


Figure 4.7 – Illustration of Ti^{4+}-IMAC adsorbent. Polystyrene microspheres are functionalized with a poly(GMA-co-TMPTMA) linker, which is coupled to phosphonate groups. These groups are able to chelate the Ti^{4+} ions used for phosphopeptide enrichment. (Adapted from [22])

The tryptic peptide mixture was loaded into a Ti^{4+}-IMAC GELoader spin tip or an Eppendorf containing biomimetic ligand-functionalized magnetic beads using 80% (v/v) ACN + 6% (v/v) TFA as loading buffer. The presence of ACN in the loading buffer reduces the binding of non-phosphorylated peptides by disrupting non-specific hydrophobic interactions, being more effective than MeOH, EtOH, and acetone [40]. The optimal amount of ACN depends on the method and is typically between 40-80% (v/v) [41-43]. TFA plays two roles in the enrichment process: (i) it efficiently

protonates acidic amino acids, preventing their non-specific binding; and (ii) it disrupts intramolecular bonds that often form between the phosphate group and basic residues. These intramolecular bonds are the cause of the higher frequency of missed cleavages during phosphoprotein digestion, as the enzyme cannot efficiently access the cleavage site [22]. The influence of different acids in phosphopeptide binding has been previously studied. Kokubu and co-workers reported that the selectivity of IMAC/C18 tips towards phosphopeptides increases with the acid strength, i.e. TFA ≥ hydrochloric acid > formic acid > acetic acid [40]. The effect of acids was also evaluated for TiO_2 tips by Aryal and co-workers, who found that the ability of organic acids to exclude non-phosphorylated peptides follows a similar order: TFA > formic acid > acetic acid [44]. The concentration of the acid plays a significant role as well, but the optimal concentration depends on the method used. Extremely low concentrations do not allow complete protonation of the acidic residues, while extremely high concentrations lead to incomplete deprotonation of the phosphate moiety [40, 44]. The addition of salt to the first washing buffer (Washing Buffer 1: 50% (v/v) ACN + 0.5% (v/v) TFA + 200 mM NaCl) assists the elimination of non-specific electrostatic adsorption. A second washing step using a buffer without salt (Washing Buffer 2: 50% (v/v) ACN + 0.1% (v/v) TFA) was performed before the elution step. Finally, the peptides were eluted using an alkaline solution of 10% (v/v) NH_4OH pH 11. The buffers have been previously optimized for Ti^{4+}-IMAC [22, 39, 41]. Figure 4.8 presents a schematic representation of the methodology developed in this work.

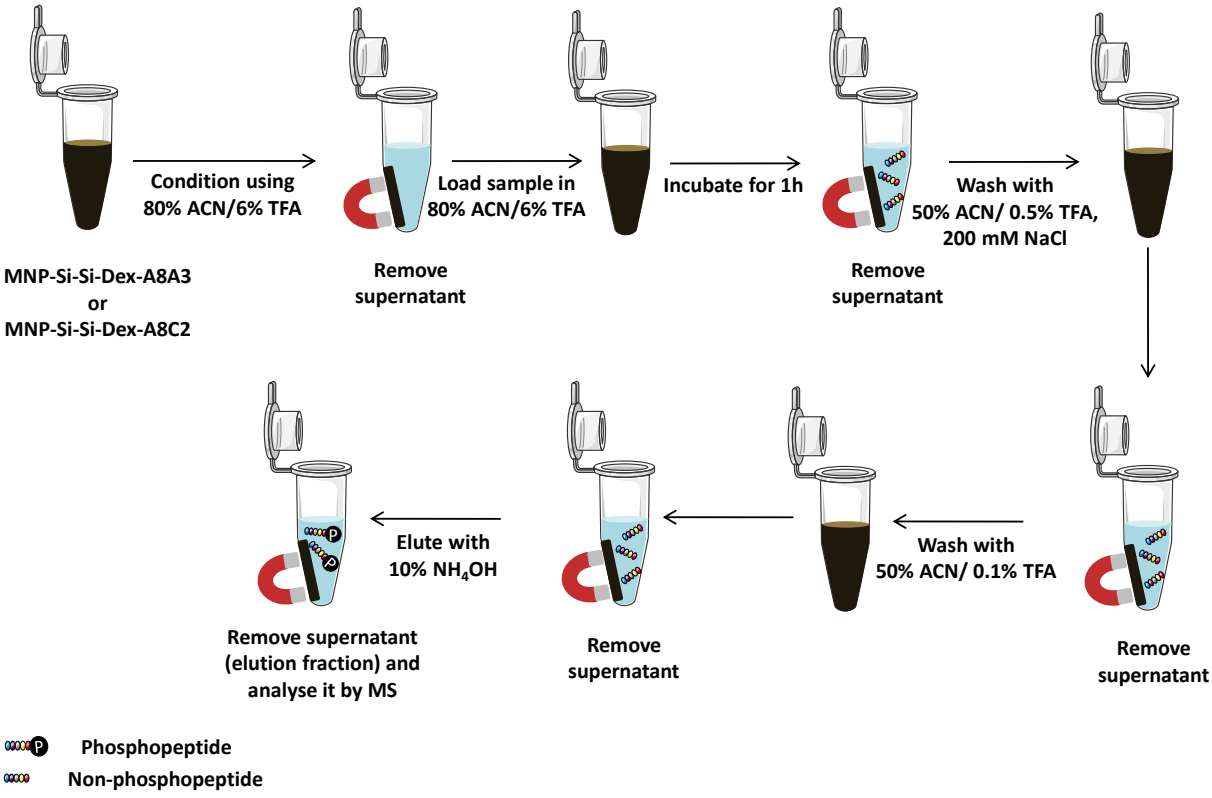


Figure 4.8 – Phosphopeptide enrichment using MNP-Si-Si-Dex functionalized with biomimetic ligands.

The eluates were dried, resuspended in MALDI matrix solution (25mg/mL 2,5-DHB in 70 % (v/v) ACN/1% (w/v) H_3PO_4) and spotted (0.5 µL) in a MALDI target plate. Several matrices (e.g. 2,5-DHB, α-cyano-4-hydroxycinnamic acid, 2,4,6-trihydroxyacetophenone), co-matrices (e.g. 1,8-bis(dimethyl-amino)naphthalene) and additives (e.g. H_3PO_4, diammonium hydrogen citrate) have been tested and optimized for phosphopeptide identification using MALDI MS, and their suitability depends on the enrichment method used and on whether the analysis is performed in positive or negative ion mode [45-47]. 2,5-DHB generates large and heterogeneous crystalline deposits, but it is considered a "cool" matrix. In other words, it induces the formation of molecular ions with low internal energy, which are stable during the MS analysis. The addition of 1% (w/v) H_3PO_4 to 2,5-DHB MALDI matrix has been reported to enhance phosphopeptide ion signals, probably due to *salting out* effects which will lead to a more efficient incorporation of the phosphopeptides into the growing crystals of the matrix [47].

MALDI-TOF MS analysis was performed for the semi-complex mixture: (i) non-enriched; (ii) enriched using Ti^{4+}-IMAC; (iii) enriched using MNP-Si-Si-Dex-A8A3; and (iv) enriched using MNP-Si-Si-Dex-A8C2 (Figure 4.9). As can be seen in Figure 4.9D, the mass spectra corresponding to the enriched sample

with MNP-Si-Si-Dex-A8C2 presents a very low signal to noise (S/N) ratio and no mass peaks corresponding to the loaded phosphopeptides were detected and identified. This might be due to low peptide concentration in the elution fraction, or to the poor co-crystallization of the sample with the MALDI matrix due to the presence of contaminants in the eluate solution. Regarding the sample processed with MNP-Si-Si-Dex-A8A3, the spectrum in Figure 4.9C is dominated by phosphorylated peptides, which shows the success of this enrichment process. Furthermore, the S/N ratio of the detected peptides is comparable to the mass spectrum of the sample enriched with Ti^{4+}-IMAC (Figure 4.9B). The list of phosphopeptides identified with the different approaches, Ti^{4+}-IMAC and MNP-Si-Si-Dex-A8A3, is presented in Table 4.2 and Table 4.3, respectively. The number of phosphopeptides identified with both methodologies is the same (9 phosphopeptides). However, no quantitative conclusion can be drawn from these experiments regarding the efficiency of both methods. It should also be kept in mind that the loading, washing, and elution conditions used here for both enrichment methods were the optimal conditions for Ti^{4+}-IMAC. The conditions for this novel system still need further optimization; therefore an improvement of the results here reported is expected.

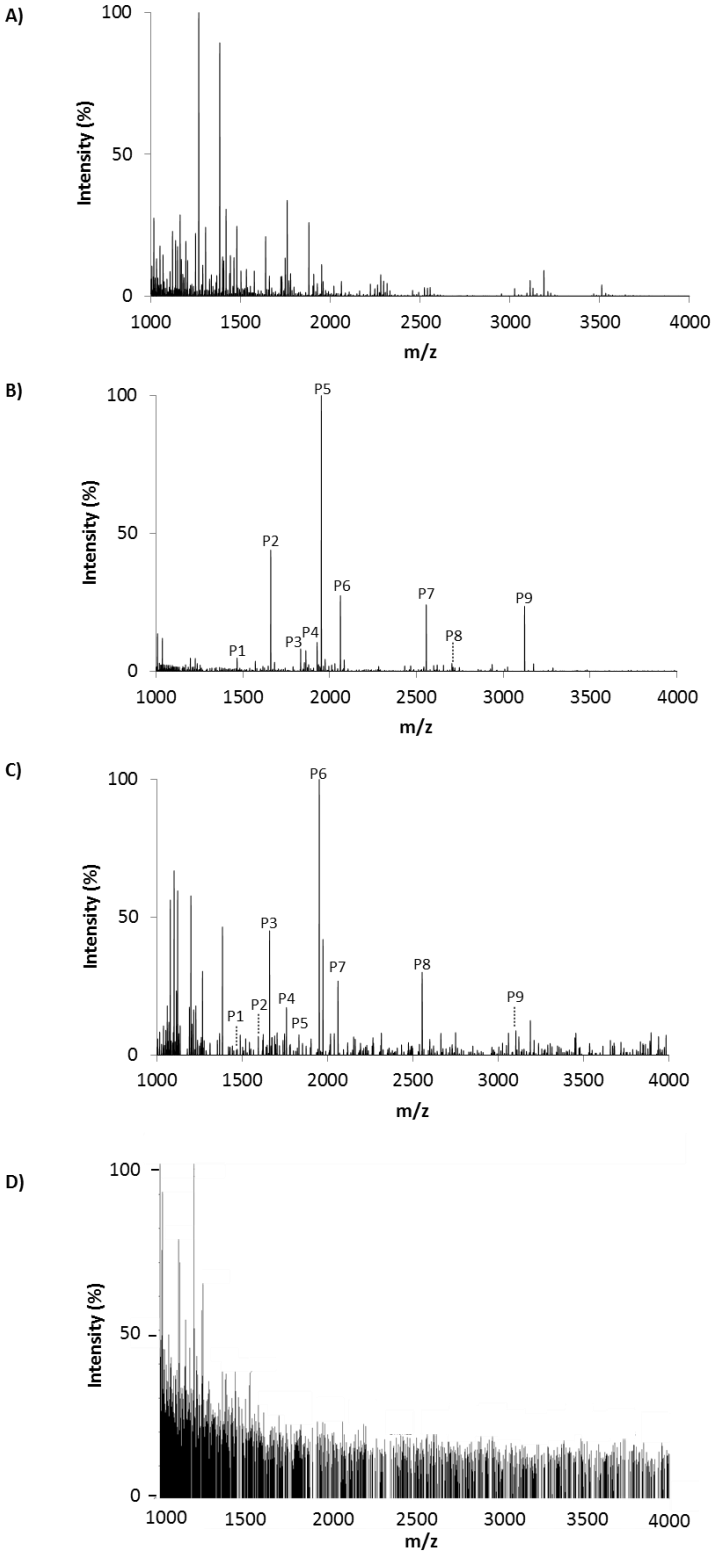


Figure 4.9 – MALDI-TOF mass spectra of a semi-complex mixture of α-casein, β-casein, and BSA: (A) before enrichment; (B) after enrichment with Ti^{4+}-IMAC; (C) after enrichment with MNP-Si-Si-Dex-A8A3; (D) after enrichment with MNP-Si-Si-Dex-A8C2. Samples were dissolved in 10 µL 25mg/mL 2,5-DHB in 70 % (v/v) ACN/1% (w/v) H_3PO_4 and 0.5 µL was spotted on the MALDI plate for analysis. Peaks correspondent to phosphopeptides are marked with letter P and the respective sequence is described in Table 4.2 and Table 4.3.

Table 4.2 – Phosphopeptides detected by MALDI-TOF MS analysis from a semi-complex mixture of α-casein, β-casein, and BSA, after enrichment with Ti^{4+}-IMAC.

Peptide ID	Peptide sequence	No. of missed cleavages	No. of phosphorylation sites	$[M+H]^+$ (Da)	Protein
P1	TVDME<u>S</u>TEVFTK	0	1	1466.61	α-s2 casein
P2	VPQLEIVPN<u>S</u>AEER	0	1	1660.77	α-s1 casein
P3	YLGEYLIVPN<u>S</u>AEER	0	1	1832.84	α-s1 casein
P4	DIG<u>S</u>E<u>S</u>TEDQAMEDIK	0	2	1927.68	α-s1 casein
P5	YKVPQLEIVPN<u>S</u>AEER	1	1	1951.94	α-s1 casein
P6	FQ<u>S</u>EEQQQTEDELQDK	0	1	2061.82	β-casein
P7	FQ<u>S</u>EEQQQTEDELQDKIHPF[a]	1	1	2556.10	β-casein
P8	Q*MEAE<u>S</u>IS<u>SS</u>EEIVPN<u>S</u>VEQK	0	5	2703.88	α-s1 casein
P9	RELEELNVPGEIVE<u>S</u>L<u>SSS</u>EESITR	1	4	3122.24	β-casein

[a] Chymotrypsin cleavage.
Q* - pyroglutamination on Gln; <u>S</u> - phosphoserine.

Table 4.3 - Phosphopeptides detected by MALDI-TOF MS analysis from a semi-complex mixture of α-casein, β-casein, and BSA, after enrichment with MNP-Si-Si-Dex-A8A3.

Peptide ID	Peptide sequence	No. of missed cleavages	No. of phosphorylation sites [b]	$[M+H]^+$ (Da)	Protein
P1	TVDME<u>S</u>TEVFTK	0	1	1466.59	α-s2 casein
P2	TVDME<u>S</u>TEVFTKK	1	1	1594.76	α-s2 casein
P3	VPQLEIVPN<u>S</u>AEER	0	1	1660.79	α-s1 casein
P4	TVD[Mo]E<u>S*</u>TEVFTKKTK	2	0 (1)	1759.93	α-s2 casein
P5	YLGEYLIVPN<u>S</u>AEER	0	1	1832.86	α-s1 casein
P6	YKVPQLEIVPN<u>S</u>AEER	1	1	1951.94	α-s1 casein
P7	FQ<u>S</u>EEQQQTEDELQDK	0	1	2061.81	β-casein
P8	KNTMEHV<u>S*S*S*</u>EE<u>S*</u>IISQETY[a]	1	3 (4)	2555.00	α-s2 casein
P8	NTMEHV<u>S*S*S*</u>EE<u>S*</u>IISQETYK	0	3 (4)	2555.00	α-s2 casein
P9	RELEELNVPGEIVE<u>S</u>L<u>SSS</u>EESITR	1	4	3122.19	β-casein

[a] Chymotrypsin cleavage.
[b] Number of phosphorylation sites if there is no hydrolysis of the phosphate group during strong alkaline elution is given in parenthesis.
[Mo] – oxidized Met; <u>S</u> – phosphoserine; <u>S</u>* - Potential phosphorylation site.

4.4. Conclusions

A novel magnetic adsorbent for the selective capture of phosphorylated peptides has been developed. Iron oxide MNPs have been coated with silica and dextran and functionalized with a triazine and an Ugi-based ligand. To our knowledge, this is the first time that small synthetic biomimetic ligands have been used for phospho-fishing. Ligand A8A3 was able to selectively capture phosphopeptides from 6 pmol of a semi-complex sample of α-casein, β-casein, and BSA. The MS spectrum presented nine peaks correspondent to phosphorylated peptides with increased signal intensities, when compared to non-enriched samples.

4.5. Bibliography

1. Veiseh, O., J.W. Gunn, and M. Zhang, *Design and fabrication of magnetic nanoparticles for targeted drug delivery and imaging.* Advanced Drug Delivery Reviews, 2010. **62**(3): p. 284-304.
2. Wu, W., Q. He, and C. Jiang, *Magnetic iron oxide nanoparticles: synthesis and surface functionalization strategies.* Nanoscale Research Letters, 2008. **3**(11): p. 397-415.
3. Laurent, S., D. Forge, M. Port, A. Roch, C. Robic, L. Vander Elst, and R.N. Muller, *Magnetic iron oxide nanoparticles: synthesis, stabilization, vectorization, physicochemical characterizations, and biological applications.* Chemical Reviews, 2008. **108**(6): p. 2064-2110.
4. McCarthy, J.R. and R. Weissleder, *Multifunctional magnetic nanoparticles for targeted imaging and therapy.* Advanced Drug Delivery Reviews, 2008. **60**(11): p. 1241-1251.
5. Huber, D.L., *Synthesis, properties, and applications of iron nanoparticles.* Small, 2005. **1**(5): p. 482-501.
6. Sun, C., J.S. Lee, and M. Zhang, *Magnetic nanoparticles in MR imaging and drug delivery.* Advanced Drug Delivery Reviews, 2008. **60**(11): p. 1252-1265.
7. Lu, A.H., E.e.L. Salabas, and F. Schüth, *Magnetic nanoparticles: synthesis, protection, functionalization, and application.* Angewandte Chemie International Edition, 2007. **46**(8): p. 1222-1244.
8. Ditsch, A., P.E. Laibinis, D.I. Wang, and T.A. Hatton, *Controlled clustering and enhanced stability of polymer-coated magnetic nanoparticles.* Langmuir, 2005. **21**(13): p. 6006-6018.
9. Lee, A., H.J. Yang, E.S. Lim, J. Kim, and Y. Kim, *Enrichment of phosphopeptides using bare magnetic particles.* Rapid Communications in Mass Spectrometry, 2008. **22**(16): p. 2561-2564.
10. Leitner, A., *Phosphopeptide enrichment using metal oxide affinity chromatography.* TrAC Trends in Analytical Chemistry, 2010. **29**(2): p. 177-185.
11. Chen, C.-T., L.-Y. Wang, and Y.-P. Ho, *Use of polyethylenimine-modified magnetic nanoparticles for highly specific enrichment of phosphopeptides for mass spectrometric analysis.* Analytical and Bioanalytical Chemistry, 2011. **399**(8): p. 2795-2806.
12. Yu, Q., X.S. Li, B.F. Yuan, and Y.Q. Feng, *Preparation of magnetic hydroxyapatite clusters and their application in the enrichment of phosphopeptides.* Journal of Separation Science, 2014.

13. Li, X.-S., L.-D. Xu, G.-T. Zhu, B.-F. Yuan, and Y.-Q. Feng, *Zirconium arsenate-modified magnetic nanoparticles: preparation, characterization and application to the enrichment of phosphopeptides.* Analyst, 2012. **137**(4): p. 959-967.
14. Sun, Y. and H.-F. Wang, *Ultrathin-yttrium phosphate-shelled polyacrylate-ferriferrous oxide magnetic microspheres for rapid and selective enrichment of phosphopeptides.* Journal of Chromatography A, 2013. **1316**: p. 62-68.
15. Zhang, L., Q. Zhao, Z. Liang, K. Yang, L. Sun, L. Zhang, and Y. Zhang, *Synthesis of adenosine functionalized metal immobilized magnetic nanoparticles for highly selective and sensitive enrichment of phosphopeptides.* Chemical Communications, 2012. **48**(50): p. 6274-6276.
16. Cheng, G., Y.-L. Liu, J.-L. Zhang, D.-H. Sun, and J.-Z. Ni, *Lanthanum silicate coated magnetic microspheres as a promising affinity material for phosphopeptide enrichment and identification.* Analytical and Bioanalytical Chemistry, 2012. **404**(3): p. 763-770.
17. Wang, Z.-G., G. Cheng, Y.-L. Liu, J.-L. Zhang, D.-H. Sun, and J.-Z. Ni, *Magnetic γ-Fe 2 O 3@ REVO 4 (RE= Sm, Dy, Ho) affinity microspheres for selective capture, fast separation and easy identification of phosphopeptides.* Journal of Materials Chemistry B, 2013. **1**(10): p. 1491-1500.
18. Zhu, L., J. Zhang, S. Ren, and Y. Guo, *Facile and fast enrichment of phosphopeptides prior to matrix-assisted laser desorption/ionization time-of-flight mass spectrometric analysis using natural nanoparticle-bentonite.* International Journal of Mass Spectrometry, 2013. **343**: p. 23-27.
19. Wang, Z.-G., G. Cheng, Y.-L. Liu, J.-L. Zhang, D.-H. Sun, and J.-Z. Ni, *Novel 3D flowerlike hierarchical γ-Fe2O3@ xNH4F· yLuF3 core–shell microspheres tailor-made by a phase transformation process for the capture of phosphopeptides.* Journal of Materials Chemistry B, 2013. **1**(37): p. 4845-4854.
20. Deng, Q., J. Wu, Y. Chen, Z. Zhang, Y. Wang, G. Fang, S. Wang, and Y. Zhang, *Guanidinium functionalized superparamagnetic silica spheres for selective enrichment of phosphopeptides and intact phosphoproteins from complex mixtures.* Journal of Materials Chemistry B, 2014.
21. Tsunehiro, M., Y. Meki, K. Matsuoka, E. Kinoshita-Kikuta, E. Kinoshita, and T. Koike, *A Phos-tag-based magnetic-bead method for rapid and selective separation of phosphorylated biomolecules.* Journal of Chromatography B, 2013. **925**: p. 86-94.
22. Zhou, H., M. Ye, J. Dong, E. Corradini, A. Cristobal, A.J. Heck, H. Zou, and S. Mohammed, *Robust phosphoproteome enrichment using monodisperse microsphere–based immobilized titanium (IV) ion affinity chromatography.* Nature Protocols, 2013. **8**(3): p. 461-480.
23. Yu, Z., G. Han, S. Sun, X. Jiang, R. Chen, F. Wang, R.a. Wu, M. Ye, and H. Zou, *Preparation of monodisperse immobilized Ti^{4+} affinity chromatography microspheres for specific enrichment of phosphopeptides.* Analytica Chimica Acta, 2009. **636**(1): p. 34-41.
24. Li, D., W.Y. Teoh, J.J. Gooding, C. Selomulya, and R. Amal, *Functionalization strategies for protease immobilization on magnetic nanoparticles.* Advanced Functional Materials, 2010. **20**(11): p. 1767-1777.
25. Migneault, I., C. Dartiguenave, M.J. Bertrand, and K.C. Waldron, *Glutaraldehyde: behavior in aqueous solution, reaction with proteins, and application to enzyme crosslinking.* Biotechniques, 2004. **37**(5): p. 790-806.
26. Masárová, J., D. Mislovicová, and P. Gemeiner, *Optimization of dextran and mannan dialdehydes preparation and examination of their biospecific interaction with concanavalin A.* Chemical Papers-Slovak Academy of Sciences, 2001. **55**(2): p. 130-135.
27. Miksa, D., E.R. Irish, D. Chen, R.J. Composto, and D.M. Eckmann, *Dextran Functionalized Surfaces via Reductive Amination: Morphology, Wetting, and Adhesion.* Biomacromolecules, 2005. **7**(2): p. 557-564.

28. *Stokes Einstein: Relating Particle Size to Particle Motion*. Available from: http://www.horiba.com/scientific/products/particle-characterization/technology/dynamic-light-scattering/, Access Date: 18 March 2014.
29. *Zetasizer Nano Series User Manual - MAN 0317 Issue 1.1 Feb. 2004.* Malvern Instruments Ltd, 2004: p. 13.1-13.2.
30. *Dynamic Light Scattering*. Available from: http://www.malvern.com/en/products/technology/dynamic-light-scattering/, Access Date: 16 March 2014.
31. Batalha, I.L., A. Hussain, and A. Roque, *Gum Arabic coated magnetic nanoparticles with affinity ligands specific for antibodies.* Journal of Molecular Recognition, 2010. **23**(5): p. 462-471.
32. Taylor, J.I., C.D. Hurst, M.J. Davies, N. Sachsinger, and I.J. Bruce, *Application of magnetite and silica–magnetite composites to the isolation of genomic DNA.* Journal of Chromatography A, 2000. **890**(1): p. 159-166.
33. Deng, Y.-H., C.-C. Wang, J.-H. Hu, W.-L. Yang, and S.-K. Fu, *Investigation of formation of silica-coated magnetite nanoparticles via sol–gel approach.* Colloids and Surfaces A: Physicochemical and Engineering Aspects, 2005. **262**(1): p. 87-93.
34. Grüttner, C. and J. Teller, *New types of silica-fortified magnetic nanoparticles as tools for molecular biology applications.* Journal of Magnetism and Magnetic Materials, 1999. **194**(1): p. 8-15.
35. Santana, S.D., V.L. Dhadge, and A.C. Roque, *Dextran-Coated Magnetic Supports Modified with a Biomimetic Ligand for IgG Purification.* ACS Applied Materials & Interfaces, 2012. **4**(11): p. 5907-5914.
36. *Zetasizer Nano Series User Manual - MAN 0317 Issue 1.1 Feb. 2004.* Malvern Instruments Ltd, 2004: p. 15.1-15.2.
37. Sun, Y., L. Duan, Z. Guo, Y. DuanMu, M. Ma, L. Xu, Y. Zhang, and N. Gu, *An improved way to prepare superparamagnetic magnetite-silica core-shell nanoparticles for possible biological application.* Journal of Magnetism and Magnetic Materials, 2005. **285**(1): p. 65-70.
38. Doane, T.L., C.-H. Chuang, R.J. Hill, and C. Burda, *Nanoparticle ζ-potentials.* Accounts of Chemical Research, 2011. **45**(3): p. 317-326.
39. Zhou, H., T.Y. Low, M.L. Hennrich, H. van der Toorn, T. Schwend, H. Zou, S. Mohammed, and A.J. Heck, *Enhancing the identification of phosphopeptides from putative basophilic kinase substrates using Ti (IV) based IMAC enrichment.* Molecular & Cellular Proteomics, 2011. **10**(10): p. M110. 006452.
40. Kokubu, M., Y. Ishihama, T. Sato, T. Nagasu, and Y. Oda, *Specificity of immobilized metal affinity-based IMAC/C18 tip enrichment of phosphopeptides for protein phosphorylation analysis.* Analytical Chemistry, 2005. **77**(16): p. 5144-5154.
41. Zhou, H., M. Ye, J. Dong, G. Han, X. Jiang, R. Wu, and H. Zou, *Specific phosphopeptide enrichment with immobilized titanium ion affinity chromatography adsorbent for phosphoproteome analysis.* Journal of Proteome Research, 2008. **7**(9): p. 3957-3967.
42. Jensen, S.S. and M.R. Larsen, *Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques.* Rapid Communications in Mass Spectrometry, 2007. **21**(22): p. 3635-3645.
43. Nabetani, T., Y.J. Kim, M. Watanabe, Y. Ohashi, H. Kamiguchi, and Y. Hirabayashi, *Improved method of phosphopeptides enrichment using biphasic phosphate-binding tag/C18 tip for versatile analysis of phosphorylation dynamics.* Proteomics, 2009. **9**(24): p. 5525-5533.
44. Aryal, U.K. and A.R. Ross, *Enrichment and analysis of phosphopeptides under different experimental conditions using titanium dioxide affinity chromatography and mass spectrometry.* Rapid Communications in Mass Spectrometry, 2010. **24**(2): p. 219-231.

45. Yang, X., H. Wu, T. Kobayashi, R.J. Solaro, and R.B. van Breemen, *Enhanced ionization of phosphorylated peptides during MALDI TOF mass spectrometry.* Analytical Chemistry, 2004. **76**(5): p. 1532-1536.
46. Zhang, S. and Z.-P. Yao, *Improved detection of phosphopeptides by negative ion matrix-assisted laser desorption/ionization mass spectrometry using a proton sponge co-matrix.* Analytica Chimica Acta, 2012. **711**: p. 77-82.
47. Kjellström, S. and O.N. Jensen, *Phosphoric acid as a matrix additive for MALDI MS analysis of phosphopeptides and phosphoproteins.* Analytical Chemistry, 2004. **76**(17): p. 5109-5117.

Chapter 4 Phosphopeptide enrichment using biomimetic magnetic nanostructures

CHAPTER 5

Development of β-hairpin peptidomimetics as phosphate-binding ligands

Summary

Cyclic β-hairpin structures have been developed by structural-based design as peptidomimetics of the Brca1 C-terminal (BRCT) domain. The amino acid residues and spatial constraints involved in the molecular recognition of a phosphorylated peptide by the BRCT domain were identified *in silico* and crafted onto a 14-mer β-hairpin template. The cyclic β-hairpins were synthesized using standard Fmoc solid-phase chemistry, purified by high-performance liquid chromatography (HPLC), and immobilized onto a chromatographic support (agarose). One of the β-hairpin structures (Cyclic-M2: Cyclo(Glu-Gly-Phe-Gly-Dap-Gly-Dap-^{D}Pro-Pro-Gly-Val-Arg-Thr-Gly), where Dap is diaminopropionic acid) presented selectivity for binding to the phosphorylated target, allowing for a recovery yield of 65% upon elution.

Chapter 5 Development of β-hairpin peptidomimetics as phosphate-binding ligands

5.1. Introduction

The term "peptidomimetics" dates from the 1980´s and its concept has been evolving in the last decades. It refers to the design and preparation of analogues of natural peptides, protein fragments, or even entire proteins [1]. Peptidomimetics are valuable tools for understanding the interplay between protein structure and function, besides their potential commercial use as therapeutic agents [2, 3]. A peptidomimetic can be a fully synthetic molecule, a peptoid (i.e. oligomer of N-substituted glycines), a β-peptide, a modified natural peptide (for example by the incorporation of non-natural amino acids), among others [4].

Research was initially focused on the discovery of small drug-like compounds, but has shifted to more complex molecules with secondary structure and conformational restrictions [2]. For *in vivo* applications, peptides are valuable alternatives to small synthetic molecules, as they present lower toxicity and higher target selectivity [5]. In particular, cyclic peptides have the advantage of being more resistant to proteolytic digestion, because they do not present terminal carboxyl and amine groups. Their conformational and structural rigidity confers them superior receptor selectivity, improved biological activity, and biochemical stability [5, 6]. Cyclic peptides usually present 10 to 1000-fold higher target-affinity when compared to their linear counterparts [7]. Cyclic peptides may be synthesized by either biological or synthetic methods. Biological methods include phage display, mRNA display, and split-intein mediated circular ligation. Synthetic methods comprise sequential synthesis, parallel synthesis, and split-and-pool synthesis [5].

Phage display allows the generation of large libraries of peptides (≈ 10^9), which are displayed at the surface of phages. These phages are screened against the target of interest, and the ones that selectively bind the target are then amplified in bacterial cells. This process is called biopanning and is usually repeated 3 to 5 times [8]. The incorporation of Cys residues in peptide sequence allows their cyclization through the formation of disulfide bonds, which will naturally occur inside the bacterial cell. However, disulfide bonds are not stable and may be reduced to form thiol groups again [5, 7]. In spite of allowing great complexity, phage display is limited to the use of the 20 natural amino acids [5].

In mRNA display, encoded peptide and protein libraries are covalently linked to the 3' end of their own mRNA through a DNA-puromycin linker. Peptides are then screened against their targets and positive hits may be amplified by reverse transcription polymerase chain reaction (RT-PCR), followed by more rounds of selection [9, 10]. This method allows constructing libraries of 10^{13} molecules, increasing sequence complexity when compared to phage display, besides allowing the incorporation

of unnatural amino acids [10]. On the other hand, it has the disadvantage of producing only one peptide per mRNA molecule, which means that if the peptide fails to couple to the mRNA template, a good binder may be lost during the selection [5]. Millward and co-workers have prepared cyclic peptides using mRNA display technology by post-translational crosslinking with disuccinimidyl glutarate of the terminal amine group of the peptide to a Lys side chain in the C-terminal region [7]. However, cyclization may not be selective if the peptide presents more amino acid residues with amine side chains besides the fixed Lys [5].

Inteins are protein sequences which catalyse their excision from precursor proteins and ligate their flanking protein sequences (called N- and C-exteins) [11, 12]. Split-intein mediated circular ligation is a technique based on the manipulation of protein splicing, which uses the *Synechocystis* sp PCC6803 *DnaE* split intein to produce libraries of cyclic peptides ($\approx 10^8$) [11, 13, 14]. This technique is again limited by the use of the natural ribosomal amino acids [5].

Synthetic peptide libraries may be constructed using sequential synthesis, i.e. the synthesis of each peptide individually from the beginning to the end. However, this process does not permit great variability and it is much time-consuming [5]. Parallel synthesis allows the simultaneous preparation of hundreds or thousands of peptides, which remains well below the values obtained with biological methods [5, 15, 16].

Millions of peptides may be synthesized using the "split-and-pool" protocol, which consists of distributing resin beads between reaction vessels containing different amino acids. After the first amino acid coupling, the beads are pooled together and then split again for another cycle [17]. The main disadvantage of this method is the difficulty to determine peptide sequences after the screening for target-binding. Edman degradation is not applicable to cyclic peptides, because they do not possess an N-terminal group. On the other hand, cyclic peptide sequencing using mass spectrometry is often difficult because ring opening occurs at multiple positions, which complicates spectral interpretation [5, 6]. A few strategies have been developed to overcome these issues. Joo and co-workers reported a spatial segregation approach consisting of segregating a resin bead into two layers: an outer layer displaying the cyclic peptide and an inner layer carrying the correspondent linear peptide. The linear peptide is used for sequencing, whereas the cyclic peptide may be screened against the target molecule [6]. Kwon and co-workers developed a different "one bead two compounds" approach relying on differential deprotection to create two very similar peptoids (although the method is also feasible for peptides). The only difference between the peptoids is that only one of them contains a Cys to promote the binding to a maleimide-activated support and a Glu residue to facilitate cyclization. In this way, only the cyclic peptide will bind to the target [18].

In summary, biological methods for peptide synthesis provide great diversity, but are usually more limited in terms of building blocks. Peptide sequence is also facilitated because of the phenotype/genotype association. Synthetic methods are more limited in quantitative terms, but still provide high complexity as they allow chemical versatility.

The most common secondary structures of proteins are α-helices and β-sheets [19, 20]. These constitute great scaffolds for biomimetic design, and libraries of peptides with both of these conformations have already been reported. Mimicking spatial orientations is fundamental when developing a peptide library as the conformational flexibility of peptides in solution has been linked to decreased binding to their targets [21]. This work will focus on the development of β-hairpin peptides, which are the simplest form of antiparallel β-sheets.

A β-hairpin is composed of two antiparallel β-strands connected by a turn/ loop [21, 22] (Figure 5.1).

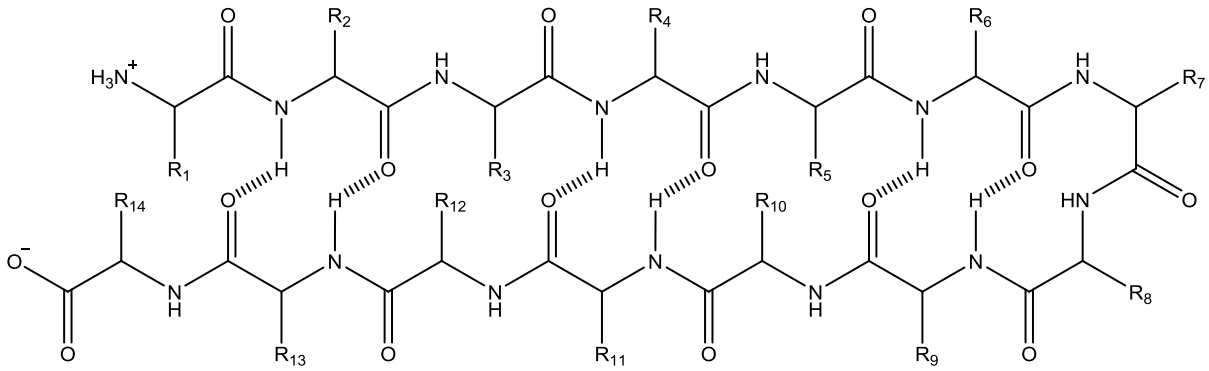


Figure 5.1 – Schematic representation of a 14-mer β-hairpin. Interchain hydrogen bonds are represented by dashed lines.

The mechanism for β-hairpin folding is not yet fully understood due to their inherent tendency to aggregate in solution, and because β-hairpins have limited ability to fold in the absence of their tertiary protein context [21, 23]. Different mechanisms have been proposed. In 2010, Lewandowska and co-workers reviewed the results of experimental studies of the structure of β-hairpin of the B1 domain of protein G, and concluded that the "broken-zipper" was the most plausible mechanism. According to this mechanism, the turn region bends due to favourable local interactions and facilitates the formation of hydrophobic interactions between nonpolar amino acids near the turn, followed by the formation of nonpolar contacts farther from the turn. Hydrogen-bonding between residues in the backbone of the peptide seem to have a role in later stages of β-hairpin folding, assisting to fix the structure. In addition, electrostatic interactions and salt-bridges across β-strands most likely contribute for β-hairpin stability. However, the prevailing structure-determining factors

according to this model are the local interactions at the turn regions and hydrophobic interactions [23]. In 2013, Makwana and co-workers studied the effect of Tyr-His and Cys-His interacting pairs in β-hairpin stability. They concluded that the electrostatic interactions between sulphur and imidazole from the Cys-His pair lead to higher structure stabilization than the aromatic interactions of the Tyr-His pair [24]. In the same year, Danelius and co-workers reported that interstrand hydrogen-bonding does have a stabilizing effect on β-hairpin peptides [25].

Cyclization is known to stabilize the folded conformation of β-hairpins [25]. Cyclic peptides have less flexible conformations and higher extent of solvation, presenting larger diffusion coefficients [26]. Cyclization may occur in different manners: (i) head-to-tail (C-terminus to N-terminus); (ii) head-to-side chain; (iii) side chain-to-tail; and (iv) side chain-to-side chain [27]. Ring closing can be attained through a variety of covalent bonds, such as amide, disulfide, lactone, ether, thioether, etc. [5]. Recently, Park and co-workers developed a "click" cyclization methodology where they have cyclized β-hairpins in different positions through azide-alkyne chemistry [28].

β-Hairpin formation is also promoted by the incorporation of centrally positioned turn-inducing structural elements, such as ^{D}Pro-^{L}Amino acid [29]. In particular, the heterochiral diproline unit - ^{D}Pro-^{L}Pro - has been extensively studied, and is known to adopt a type II' β-turn which induces β-hairpin formation and promotes stability [29, 30].

Two faces can be defined on this type of cyclic β-hairpin structures: a binding face containing key amino acids for target recognition, and a property-tuning face comprising amino acids which are incorporated to enhance solubility and avoid aggregation (Figure 5.2) [31].

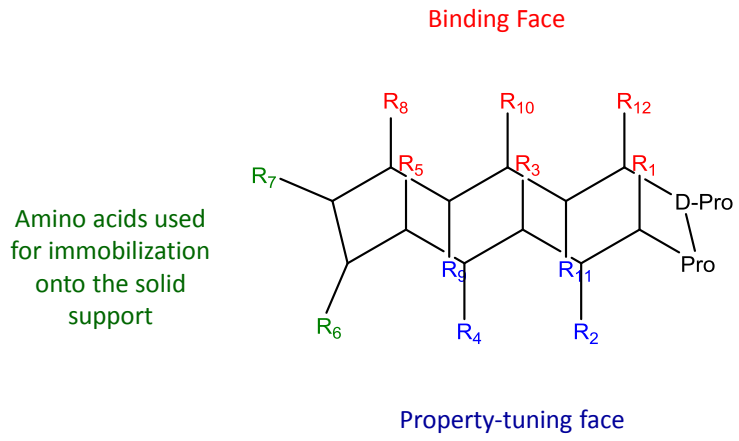


Figure 5.2 – Illustration of a ^{D}Pro-^{L}Pro β-hairpin. Amino acids comprising the binding face for target recognition are represented in red; amino acids comprising the property-tuning face to control solubility and aggregation are represented in blue; and amino acids used for attachment to the solid-support are depicted in green.

Here, it is reported the chemical synthesis of cyclized β-hairpin peptides which were designed to mimic the phosphate-binding region of the Brca1 C-terminal (BRCT) phosphoprotein-binding domain. The peptides were synthesized using Fmoc solid-phase chemistry, which allowed the introduction of chemical diversity by the use of unnatural amino acids. ^{D}Pro-^{L}Pro unit was introduced to facilitate β-hairpin formation, and peptides were cyclized through amide bonds in a head-to-tail manner. The ability of these β-hairpin peptidomimetics to bind to their phosphorylated target was also investigated.

5.2. Materials and Methods

5.2.1. Chemicals

All reagents were of the highest purity available and the solvents were HPLC-gradient.

2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBOP), H-Gly-2-ClTrt resin, and H-Glu(OtBu)-2-ClTrt resin were acquired from Novabiochem. 2-(*N*-Morpholino)ethanesulfonic acid (MES), 4-(2-hydroxyethyl)piperazine-1-ethanesulfonic acid (HEPES), 5,5'-dithiobis(2-nitrobenzoic acid) (Ellman's reagent), acetic anhydride, ammonium hydroxide solution, anisole, bicinchoninic acid kit for protein determination (BCA), diethyl ether, epichlorohydrin, ethanothiol, ethylenediaminetetraacetic acid (EDTA), ethylenediaminetetraacetic acid tetrasodium salt hydrate (EDTA-Na), iodoacetic anhydride, *N*-(3-dimethylaminopropyl)-*N*'-ethylcarbodiimide (EDC), *N,N*-diisopropylethylamine (DIEA), *N*-Hydroxysuccinimide (NHS), ninhydrin, phenol, piperidine, potassium cyanide, pyridine, sodium phosphate dibasic heptahydrate, sodium phosphate monobasic monohydrate thioanisole, trifluoroacetic acid (TFA), triisopropylsilane (TIS), tris(hydroxymethyl)aminomethane (Tris), and β-mercaptoethanol were obtained from Sigma-Aldrich. Acetonitrile (ACN), bond-Breaker tris(2-carboxyethyl)phosphine (TCEP) solution, chloroform, dichloromethane (DCM), methanol (MeOH), *N,N*-dimethylformamide (DMF), and *N*-methyl-2-pyrrolidone (NMP) were obtained from Fisher Scientific. Ethanol, sodium chloride (NaCl), and sodium hydroxide (NaOH) were purchased from Panreac. 2-Propanol, L-cysteine, and *O*-(7-azabenzotriazole-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HATU) were purchased from Roth, Acros Organics, and GenScript Corporation, respectively. Nitrogen gas was supplied by Air Liquide. N-fluorenylmethoxycarbonyl (Fmoc)-protected amino acids (Fmoc-Arg(Pbf)-OH, Fmoc-Cys(Trt)-OH,

Fmoc-Dpr(Boc)-OH (Dap), Fmoc-D-Pro-OH, Fmoc-Gly-OH, Fmoc-Ile-OH, Fmoc-Phe-OH, Fmoc-Pro-OH, Fmoc-Thr(tBu)-OH, Fmoc-Val-OH) were purchased either from CEM or Novabiochem. Gly-Ala-Ala-Tyr-Asp-Ile-Ser-Gln-Val-Phe-Pro-Phe-Ala-Lys (GK14) and Gly-Ala-Ala-Tyr-Asp-Ile-pSer-Gln-Val-Phe-Pro-Phe-Ala-Lys (GK14-P) peptides (pSer is phosphoserine) were > 98% pure and were obtained from Genecust. Unless otherwise noticed peptides contain free N- and C-terminals.

5.2.2. Chromatographic materials

Sepharose™ CL-6B was acquired from GE Healthcare. Captiva 96-well 20u filter plate with the respective duo seal 96-well Pl seal and Captiva 96-well plate cover were purchased from Agilent Technologies. Half-area UV-Star® 96-well microplates and 96-well transparent microplates were obtained from Greiner Bio-One and Sarstedt, respectively.

5.2.3. Software

Peptides were sketched on PyMOL Molecular Graphics System v.1.3 (Schrödinger), and their properties were determined using isoelectric point, logD, and pKa plugins from MarvinSketch 4.1.13 software (ChemAxon). Automated docking was performed in Autodock 4.0 software package[6].

5.2.4. Methods

5.2.4.1. Solid-phase synthesis of protected peptidomimetic scaffolds

Three fully protected linear peptides were prepared on a Liberty automated peptide synthesizer (CEM) using standard Fmoc solid-phase chemistry [32]. The sequence of the peptides was the following: Cys-Gly-Gly-Gly-Gly-Gly-Gly-^{D}Pro-Pro-Gly-Val-Gly-Thr-Gly (M0); Cys-Gly-Phe-Gly-Ile-Gly-Ile-^{D}Pro-Pro-Arg-Val-Gly-Thr-Gly (M1); Glu-Gly-Phe-Gly-Dap-Gly-Dap-^{D}Pro-Pro-Gly-Val-Arg-Thr-Gly (M2). The resins used were H-Gly-2-ClTrt for peptides M0 and M1, and H-Glu(OtBu)-2-ClTrt for peptide M2. Resin substitutions were 0.46 m_{eq}/g and 0.57 m_{eq}/g, respectively. These resins were used to obtain fully protected peptides with carboxylate end terminals after cleavage from the resin using mild acidic conditions. The activating agent was HBTU and the activator base was DIEA in NMP. 20% Piperidine in DMF was used for deprotection of the Fmoc group. After each amino acid coupling, the

[6] Automated docking was performed in our group by Dr. Ricardo Branco.

resin was end-capped with 20% acetic anhydride in DMF to block unreacted active sites. In the end, resins were thoroughly washed in a sinter funnel under vacuum pressure with DMF and DCM, and dried under vacuum.

A small aliquot of each resin was treated as indicated below to cleave and fully deprotect the peptides for electrospray ionization mass spectrometry (ESI-MS) analysis. This step was carried out to confirm if the synthesized peptides have the expected molecular weight before proceeding with cyclization. Peptides M0 and M1, containing Cys residues, were dissolved in 2 mL of 90% (v/v) TFA/ 5% (v/v) thioanisole/ 3% (v/v) ethanothiol/ 2% (v/v) anisol and peptide M2 in 2 mL of 95% TFA/ 2.5% MilliQ water/ 2.5% TIS (v/v), and stirred for 2 h under nitrogen gas. Each resin was filtered in a sinter funnel under vacuum pressure and washed with TFA. Nitrogen gas was bubbled into the filtrate in order to remove TFA. Cold diethyl ether was added to precipitate the peptide, but peptide precipitation was unsuccessful. Diethyl ether was evaporated with nitrogen gas, and an oil/ gel was obtained. This step was repeated in order to wash the crude peptide. The oily sample was then dissolved in (i) 50% MeOH (v/v) + 0.1% acetic acid for M0 and M1, or (ii) ACN for M2; and then analysed by ESI-MS using an API-linear ion trap in positive ion mode[7].

After confirming the molecular weight, the remaining resin-coupled peptides were cleaved from the resin using mild acidic conditions to obtain fully protected derivatives. The following protocol was employed: (i) 1 g of dry resin was pre-swelled in DCM in a sealable glass funnel, and then excess DCM was removed; (ii) a 10 mL solution of 1% TFA in DCM (v/v) was added to the resin, the funnel was sealed and the mixture was stirred for 2 min; and (iii) the solution was then filtered under vacuum pressure into a vial containing 800 µL of DIEA to neutralize the solution and stop the reaction. Steps (ii) and (iii) were repeated 4 times; the resin was washed with 10 mL of DCM; and steps (ii) and (iii) were repeated four additional times. Finally, the resin was washed 3 times with 10 mL of DCM. A small aliquot of each of the filtrate samples was analysed by thin layer chromatography (TLC) using a mixture of 10% MeOH in chloroform as the mobile phase. All the filtrates containing peptides were pooled together and DCM was evaporated in the rotary evaporator. dd Water was added in order to precipitate the protected peptide, but a gel was formed instead of a white precipitate. Water was then removed in the rotary evaporator.

[7] MS assays were performed in the Mass Spectrometry Laboratory at the Analytical Services Unit of the Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa.

5.2.4.2. Cyclization of M2 peptide

M2 peptide was solubilized in 5 mL of MeOH. The cyclization reaction consisted in adding aliquots of 1 mL peptide solution, followed by 30 µL DIEA (12 molar equivalents (eq.) relative to peptide) and 200 µL HATU in DMF (1.5 molar eq. HATU relative to peptide), to 25 mL DCM every 30 min under a N_2 atmosphere. During the cyclization reaction small aliquots were dried in a Schlenk line, dissolved in 50% solvent A/ 50% solvent B (v/v), where solvent A is 99.9% MilliQ water/ 0.1% TFA (v/v) and solvent B is 90% ACN/ 9.9% MilliQ water/ 0.1% TFA (v/v/v), and analysed by analytical reverse-phase HPLC (Phenomenex Jupiter Proteo column, 250 mm × 4.6 mm, 4 µm, 90 Å). The cyclic peptide solution was placed in the rotary evaporator in order to evaporate DCM, DMF, and MeOH. The cyclic peptide was then resuspended in DCM and a liquid-liquid extraction was performed with water in order to eliminate the presence of the coupling reagents and DIEA, which will be soluble in the aqueous phase. The organic phase was collected and DCM was evaporated in the rotary evaporator. Cyclic-M2 was then deprotected by solubilizing the peptide in 10 mL 95% TFA/ 2.5% MilliQ water/ 2.5% TIS (v/v) and using the same strategy used for the linear protected M2 described in §5.2.4.1. Unprotected Cyclic-M2 was then purified by preparative reverse phase HPLC (Phenomenex Jupiter Proteo column, 250 mm × 21.20 mm, 4 µm, 90 Å) using solvent A and solvent B. The peptide was eluted from the column with a linear gradient from 20% to 40% B in 30 min at a flow rate of 10 mL/min (Retention time (Rt) = 12 min). A sample of Cyclic-M2 was then analysed by Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS). The sample was dissolved in 50% (v/v) ACN and spotted onto the MALDI plate using 5 mg/mL α-CHCA in 50% (v/v) ACN/ 5% (v/v) formic acid. Mass spectrum has been acquired in the positive reflectron mode using a 4800 Plus MALDI-TOF/TOF MS analyser. Cyclic-M2 was liophilized and stored at -20°C.

5.2.4.3. Cyclization of M0 and M1 peptides

The linear protected peptide (M0 or M1) was dissolved in 20 mL MeOH, and added to a round-bottom flask containing 500 mL DCM, 3 molar eq. of PyBOP, and 4 molar eq. of DIEA. Cyclization reaction was followed by analytical reverse-phase HPLC (Phenomenex Jupiter Proteo column, 250 mm × 4.6 mm, 4 µm, 90 Å) using the same solvent system as described before. DCM and MeOH were evaporated in the rotary evaporator; the peptide was resuspended in DCM and a liquid-liquid extraction was performed with dd water. The organic phase was collected and DCM was evaporated in the rotary evaporator. Full deprotection of Cyclic-M0 and Cyclic-M1 was performed by solubilizing

the peptides in 10 mL 90% (v/v) TFA/ 5% (v/v) thioanisole/ 3% (v/v) ethanothiol/ 2% (v/v) anisol and following the procedure described in §5.2.4.1. Unprotected cyclic peptides were purified by preparative reverse phase HPLC (Phenomenex Jupiter Proteo column, 250 mm × 21.20 mm, 4 µm, 90 Å) using solvent A and solvent B. Cyclic-M0 was eluted from the column with a linear gradient from 10% to 30% B in 30 min at a flow rate of 10 mL/min (Rt = 24 min); and Cyclic-M1 was eluted from the column with a linear gradient from 20% to 35% B in 20 min at a flow rate of 10 mL/min (Rt = 11 min). Cyclic peptides were then analysed by MS. Cyclic-M0 and Cyclic-M1 were dissolved in MeOH and analysed by ESI-MS in an API-linear ion trap in positive ion mode. Both peptides were liophilized and stored at -20°C.

5.2.4.4. Ellman's test to determine free sulfhydryl groups on Cyclic-M0 and Cyclic-M1

Cyclic peptides were dissolved in reaction buffer (0.1 M sodium phosphate buffer, 1 M EDTA, pH 8) to a final concentration of 1 mM. A solution of 4 mg/mL of Ellman's reagent in reaction buffer was freshly prepared. 250 µL of the 1 mM peptide solution was added to 50 µL Ellman's reagent solution and 2.5 mL reaction buffer; the mixture was incubated at room temperature for 15 min, and the absorbance was measured at 412 nm in an EnVision Multilabel plate reader from PerkinElmer. A calibration curve was represented by standard solutions of Cys (0-1.5 mM).

5.2.4.5. Modification of agarose with iodoacetyl groups

Aminated agarose was prepared according to §2.2.4.4, and then resuspended in a solution of iodoacetic anhydride (3 molar eq. relative to epoxy groups in epoxy-activated agarose) in DMF to a final concentration of 1 g aminated agarose/ mL. DIEA was added and the mixture was incubated 2 h at room temperature with agitation (250 rpm). Finally, iodoacetylated agarose was washed with 2x DMF, 2x 50% (v/v) DMF/ dd water, and 5x dd water. The success of reaction was confirmed by the Kaiser test described in §2.2.4.4.

5.2.4.6. Immobilization of Cyclic-M0 and Cyclic-M1 onto iodoacetylated agarose

Iodoacetylated agarose was thoroughly washed with coupling buffer (50 mM Tris, 5 mM EDTA-Na, pH 8.5) in a sinter funnel under vacuum pressure. Cyclic-M0 and Cyclic-M1 (1 molar eq. relative to epoxy groups on epoxy-activated agarose) were solubilized in 1 mL of coupling buffer and 50 µL of

TCEP. The mixture was incubated at room temperature for 30 min with agitation, and an aliquot of 50 µL was taken to determine coupling efficiency. Peptide solution (1 mL) was added to 1 g of iodoacetylated agarose and incubated 15 min with agitation at room temperature, followed by additional 30 min incubation without agitation. Flow-through was collected and the resin was washed with 1 M NaCl (3x 1 mL) and coupling buffer (2x 1 mL). 1 mL of 50 mM of a blocking agent (cysteine, ethanethiol, or β-mercaptoethanol) was added to the resin, followed by 15 min incubation with agitation and 30 min incubation without agitation at room temperature. Finally, the resin was washed with coupling buffer (3x 1 mL). Absorbance of loading, flow-through and washes was measured in a microplate reader (Tecan F200) using a 280 (5) nm filter (Tecan). Calibration curves were obtained for each peptide using a 0-1 mg/mL concentration range.

5.2.4.7. Immobilization of Cyclic-M2 onto aminated agarose

Cyclic-M2 (1.5 molar eq. relative to epoxy groups in epoxy-activated agarose) was dissolved in 1 mL of activation buffer (0.1 M MES, 0.5 M NaCl, pH 6), followed by the addition of EDC (10 molar eq. relative to peptide) and NHS (to a final concentration of 5 mM) and incubation for 15 min at room temperature with mild agitation. The peptide solution was added to aminated agarose (prepared according to §2.2.4.4) and 1 M HEPES pH 8 was added until the pH of the mixture rose to 7. The mixture was incubated 2 h at room temperature with mild agitation. The resin was thoroughly washed with activation buffer. Loading, flow-through and washes were quantified using the BCA assay described in §2.2.4.13.

5.2.4.8. Screening of cyclic peptides against GK14 and GK14-P

Cyclic-M0 and Cyclic-M1 peptides were distributed between the wells of a 96-well filter plate (0.25 g of cyclic peptide-functionalized resin/well). The plates were washed with regeneration buffer (0.1 M NaOH in 30% isopropanol) followed by dd water (3 cycles of washes, 0.75 mL/well), followed by binding buffer (8x 0.75 mL/well on average). GK14 and GK14-P peptides were solubilized in binding buffer (0.5 mg/mL) and 0.25 mL of each peptide was loaded per well. The cyclic peptide-functionalized resins were incubated with GK14 and GK14-P peptides for 1h at room temperature with manual shaking. After filtration, the resins were washed with binding buffer (8x 0.25 mL). Both the flow-through and washes were collected in 96-well microplates by centrifugation at 600 rpm for 20 s, and quantified using the BCA assay described in §2.2.4.13. The same procedure was used to

Cyclic-M2, but 0.125 g of cyclic peptide-functionalized resin/well and solutions of 1 mg/mL of GK14 and GK14-P were used instead. This experiment was also performed for blank agarose, which was used as control. Binding buffers were: (i) 50 mM Gly pH 3 and 50 mM HEPES pH 8. GK14 and GK14-P were further eluted from the Cyclic-M2-functionalized resin using 50 mM HEPES, 250 mM NaCl, pH 8. Elution fractions were also quantified using BCA assay.

5.3. Results and Discussion

5.3.1. Rational design of β-hairpin cyclic peptides [8]

As previously referred in Chapter 1, the BRCT domain is a human phosphoprotein-binding domain linked with the occurrence of breast and ovarian cancers. These domains recognize the consensus sequence pSer-X-X-Phe, where X is any amino acid. The superimposition of nine crystallographic structures of the BRCA1-BRCT domain in complex with different phosphopeptide sequences confirmed the existence of a highly conserved binding mode (Figure 5.3). pSer is recognized by three highly conserved residues: Ser1655, Gly1656, and Lys1702; while Phe is stabilized by hydrogen-bonding and salt-bridging to Arg1699 and non-polar interactions with Met1775. The determinants for phosphopeptide recognition are described in detail in §1.2.1.2.

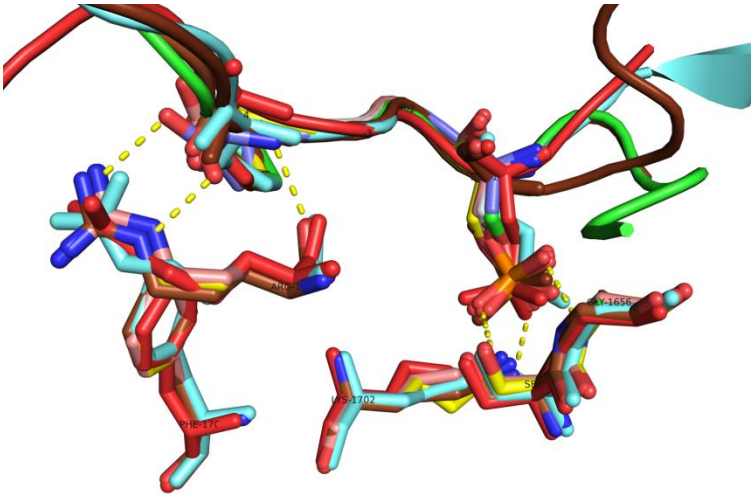


Figure 5.3 - Superposition of 9 crystallographic structures of the BRCA1-BRCT domain (PDB codes: 1T15, 1T29, 1T2V, 1Y98, 3COJ, 3K0H, 3K0K, 3K15 and 3K16).

[8] Molecular modelling studies were conducted in our laboratory by Dr. Ricardo Branco.

Chapter 5 Development of β-hairpin peptidomimetics as phosphate-binding ligands

A solution NMR structure of a 14-mer β-hairpin cyclic peptide (PDB code: 2NS4) was used as template to design a small library of peptidomimetics. The idea was to build a smaller functional mimic of the BRCT domain. The template had the following amino acid sequence: cyclo($Lys_{(1)}$-$Gly_{(2)}$-$Arg_{(3)}$-$Arg_{(4)}$-$Ile_{(5)}$-$Arg_{(6)}$-$Ile_{(7)}$-$^{D}Pro_{(8)}$-$Pro_{(9)}$-$Arg_{(10)}$-$Val_{(11)}$-$Arg_{(12)}$-$Thr_{(13)}$-$Arg_{(14)}$). Nine *in silico* point mutations were introduced in order to provide structural and functional diversity at specific positions of interest. Specifically, four Arg residues at positions 4, 6, 10 or 12, and 14 were knocked out by a Gly mutation. Arg at position 10 or 12 was conserved to prompt electrostatic interactions and hydrogen-bonding at the hydrophilic pocket, which are mainly established in the crystallographic structure of BRCA1-BRCT domain between the backbone atoms of Ser1655 and Gly1656, as well as the side chain of Lys 1702, and pSer moiety. Lys at position 1 was considered the preferential site for peptide immobilization on the solid support, as it is positioned at the opposite site of the ^{D}Pro-Pro unit which induces a β-turn on the peptide. Lys was mutated to a Gln or a Met residue which account for the anchoring point[9]. Arg at position 3 was mutated to a Phe which mimics the Phe1704 of the hydrophobic-binding pocked that accommodates the Phe of the phosphorylated peptide. Two Ile residues at positions 5 and 7 were mutated to the non-natural amino acid diaminopropionic acid (Dap), in order to mimic the side chain interactions between Lys1702 and the phosphate group of the target peptide at different interatomic distances. A control peptide (Cyclic-M0) in which all the previously mutated positions were converted to Gly residues (with the exception of the anchoring point) was also designed. Cyclic peptides were sketched in PyMOL Molecular Graphics System v.1.3 (Schrödinger) and docked with GK14-P and its non-phosphorylated version GK14 using the Autodock v.4.2 software package. GK14-P is the phosphorylated peptide which is in complex with the BRCA-BRCT domain in the crystallographic structure 1T2V (PDB code) presented in Figure 5.3.

Two β-hairpin peptides (Cyclic-M1 and Cyclic-M2) were selected as the best candidates based on the estimated docking free energy of binding. Sequences of cyclic β-hairpin peptides and their targets are presented in Table 5.1. Their affinity constants (K_a, M^{-1}) relative to the phosphorylated target GK14-P are presented in Table 5.2.

[9] Peptides were chemically synthesized using either a Glu or Cys residue for anchoring, which after immobilization to the solid-support become Gln and Met, respectively.

Table 5.1 – Amino acid sequences of cyclic β-hairpin peptides and their targets. Residues used for the immobilization to the solid-support are represented in green. Residues involved in molecular recognition are represented in red. pSer is depicted in bold.

Peptide ID	Amino acid sequence
Cyclic-M0	Cyclo(**Cys**-Gly-Gly-Gly-Gly-Gly-Gly-^{D}Pro-Pro-Gly-Val-Gly-Thr-Gly)
Cyclic-M1	Cyclo(**Cys**-Gly-**Phe**-Gly-Ile-Gly-Ile-^{D}Pro-Pro-**Arg**-Val-Gly-Thr-Gly)
Cyclic-M2	Cyclo(**Glu**-Gly-**Phe**-Gly-**Dap**-Gly-**Dap**-^{D}Pro-Pro-Gly-Val-**Arg**-Thr-Gly)
GK14	Gly-Ala-Ala-Tyr-Asp-Ile-Ser-Gln-Val-Phe-Pro-Phe-Ala-Lys
GK14-P	Gly-Ala-Ala-Tyr-Asp-Ile-**pSer**-Gln-Val-Phe-Pro-Phe-Ala-Lys

Table 5.2 – Affinity constants (Ka, M-1) of selected cyclic peptides towards the phosphorylated target GK14-P and its non-phosphorylated version.

K_a (M^{-1})		Target peptide	
		Non-Phosphorylated (GK14)	Phosphorylated (GK14-P)
β-Hairpin peptidomimetics	Cyclic-M0	3.05×10^{2}	4.85×10^{1}
	Cyclic-M1	5.93×10^{-1}	2.39×10^{3}
	Cyclic-M2	1.62×10^{-2}	3.78×10^{3}

The highest scored docking solutions of Cyclic-M1 and Cyclic-M2 in complex with the phosphorylated peptide GK14-P are depicted in Figure 5.4. The phosphate group coordinates to Arg at position 10 or 12, both positioned at the same face of the β-hairpin structure, together with a Phe at position 3.

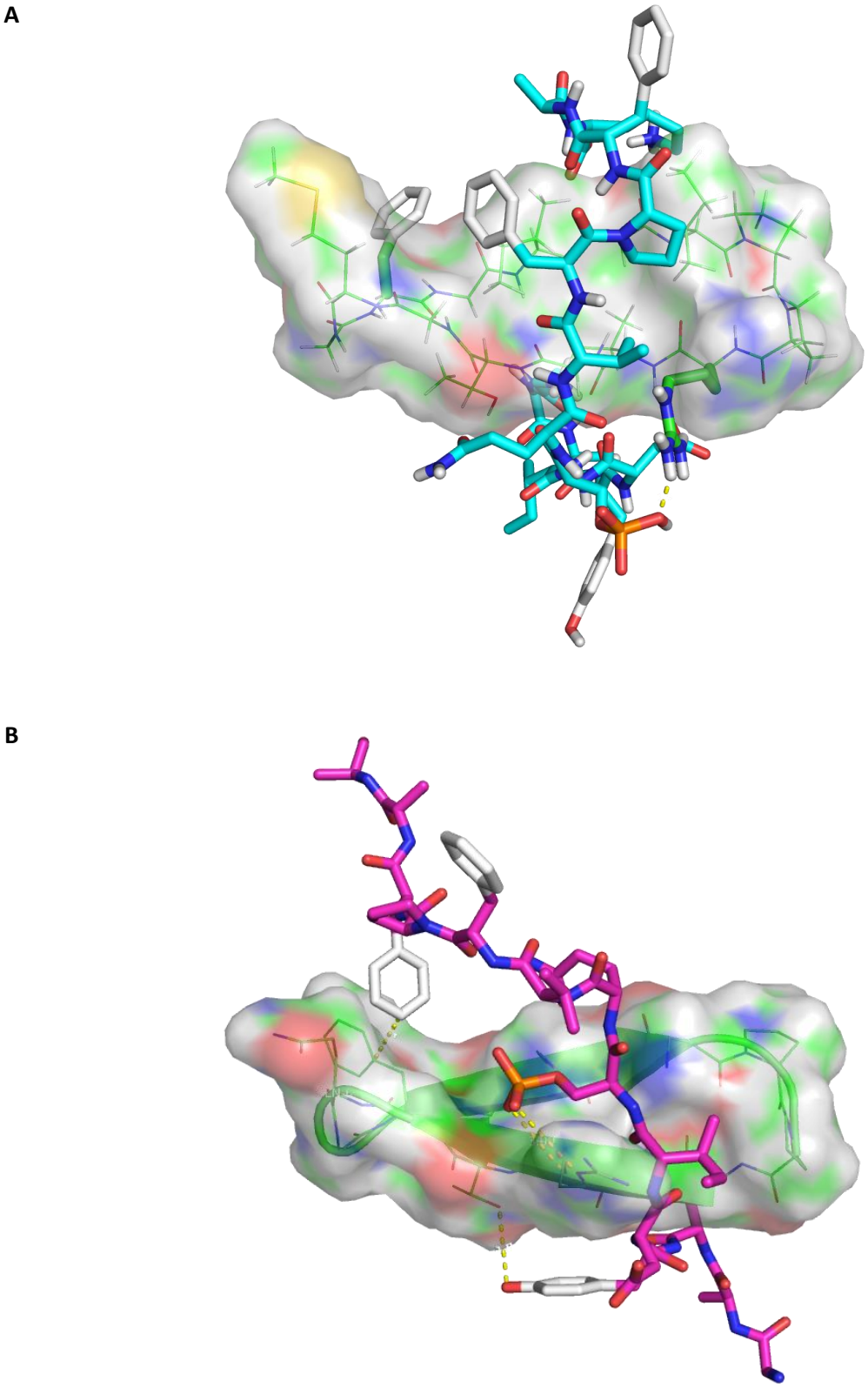


Figure 5.4 – Highest scored docking solution of A) Cyclic-M1 and B) Cyclic-M2 in complex with phosphorylated peptide GK14-P.

5.3.2. Synthesis of β-hairpin cyclic peptides

Standard microwave-assisted solid-phase Fmoc chemistry was used to prepare protected linear peptides M0, M1 and M2. 2-Chlorotrityl resins (pre-loaded with Gly or Glu) were used to obtain free carboxylic C-terminal groups. Peptides can be cyclized on the solid support, which has the advantage of facilitating the purification process. However, functional groups on solid-supported peptides are less prone to react together when compared to functional groups of peptides that freely diffuse in solution [27]. Therefore, in this work protected linear peptides were cleaved from the solid support before cyclization. Short treatments with 1% TFA (v/v) in DCM were used to cleave from the resin the fully protected peptides with free C- and N- terminal groups.

Small aliquots were fully deprotected with 90% (v/v) TFA/ 5% (v/v) thioanisole/ 3% (v/v) ethanothiol/ 2% (v/v) anisol (for M0 and M1) and 95% TFA/ 2.5% MilliQ water/ 2.5% TIS (v/v) (for M2) and analysed by ESI-MS. The mass spectrum of M0 in positive ion mode presented two high intensity peaks: *m/z* 1030 Da (z=+1) and *m/z* 1068 Da (z=+1). The first peak corresponds to the molecular mass of linear M0, and the second peak to a potassium adduct of *m/z* 1030 Da. The mass spectrum of M1 in positive ion mode presented a set of three most abundant peaks in different forms of ionization: *m/z* 1331 Da (z=+1), *m/z* 666 Da (z=+2), and *m/z* 685 Da (z=+2). The first and second peaks correspond to the molecular mass of linear M1, and the third peak to a potassium adduct of *m/z* 666 Da. The mass spectrum of M2 in positive ion mode presents a set of two most abundant peaks in different forms of ionization: *m/z* 1302 Da (z=+1) and *m/z* 652 Da (z=+2), which correspond to the molecular mass of linear M2.

Cyclization reaction of the protected peptide was performed manually in solution. The M2 peptide was cyclized using HATU as coupling reagent. HATU has been reported to give better reaction yields than other coupling reagents, such as HBTU and PyBOP [33]. A stepwise addition protocol was followed in order to avoid the use of high amounts of solvents, because cyclization reactions require the use of submillimolar peptide concentrations to avoid peptide oligo- and polymerization [27, 34]. However, in this work poor results were obtained using this approach. The cyclization reaction was followed by analytical reverse-phase HPLC and the chromatograms of Cyclic-M2 presented numerous peaks. This might be due to: (i) formation of oligomers; or (ii) presence of contaminants, due either to the fact that the linear peptide could not be precipitated in previous experimental steps and therefore could not be properly washed or to the fact that HATU reacted also with the N-terminal amine in addition to activate the C-terminal carboxylic group. Therefore, a less reactive coupling reagent that is unable to form the N-terminal adduct was used cyclize M0 and M1 peptides. HATU

was replaced by PyBOP and a larger amount of solvent was employed (to keep high dilution conditions). Indeed, the chromatograms obtained when using the latter strategy were cleaner, presenting sharp peaks correspondent to the cyclic peptides.

After cyclization, peptides were fully deprotected using the methodology described above, purified in a preparative reverse-phase HPLC, and analysed by ESI-MS in positive ion mode to confirm their molecular weight. The mass spectrum of Cyclic-M0 presented three high intensity peaks: *m/z* 1011 Da (z=+1), *m/z* 1033 Da (z=+1), and *m/z* 1049 Da (z=+1). The first peak corresponds to the molecular mass of Cyclic-M0, the second and third peaks to sodium and potassium adducts of *m/z* 1011 Da, respectively. The mass spectrum of Cyclic-M1 presented one high intensity peak at *m/z* 1312 Da (z=+1), correspondent to the molecular mass of Cyclic-M1. Cyclic-M2 was analysed by MALDI-TOF-MS and its mass spectrum presented a high intensity peak at *m/z* 1284 Da correspondent to the molecular weight of the peptide.

Cyclic peptides were then analysed by analytical reverse-phase HPLC to determine their purity. All peptides were > 95% pure.

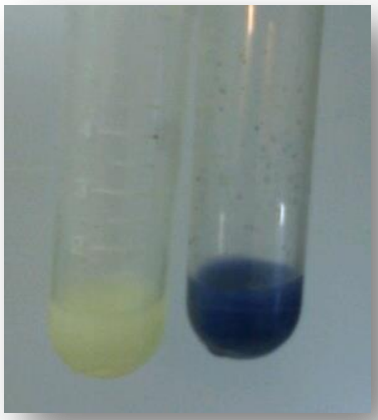
5.3.3. Immobilization of β-hairpin cyclic peptides

Two different approaches were used to immobilize β-hairpin cyclic peptides on a chromatographic support (agarose). Cyclic-M0 and Cyclic-M1 were immobilized on iodoacetylated agarose through the sulfhydryl group of their Cys residues, while Cyclic-M2 was immobilized on aminated agarose through the side chain carboxyl group of its Glu residue using carbodiimide chemistry.

The functionalization of solid supports with iodoacetyl groups has already been extensively reported [35, 36]. In this work, aminated agarose was reacted with iodoacetic anhydride to yield iodoacetyl-functionalized agarose beads. The success of the reaction was confirmed by the Kaiser test (Figure 5.5A).

Ellman's test was perfomed to determine the percentage of free sulfhydryl groups available in the peptides to react with the iodoacetylated agarose. Ellman's test is a colorimetric test based on the reaction between 5,5'-dithio-bis-(2-nitrobenzoic acid) and sulfhydryl groups, which yields a mixed disulfide and 2-nitro-5-thiobenzoate ion (TNB^{2-}). TNB^{2-} presents an intense yellow colour and absorbance at 412 nm (Figure 5.5B). Cyclic-M0 and Cyclic-M1 presented 79% and 60% of free sulfhydryl groups, respectively.

A



B

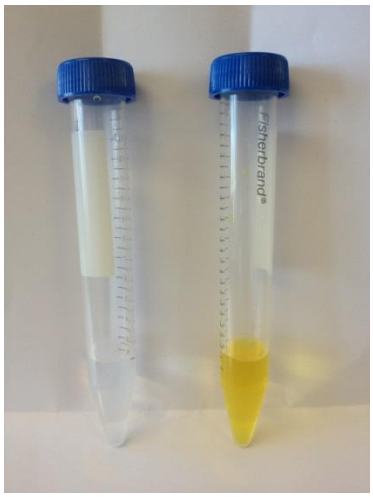


Figure 5.5 – A) Kaiser test of iodoacetyl-functionalized agarose (on the left) and aminated agarose (on the right). Aminated agarose presented a purple/blue colour characteristic of the presence of primary amines, in contrast with iodoacetylated agarose which presented a whitish colour. B) Ellman's test of 0 mM Cys (control) and 15 mM Cys in 0.1 M sodium phosphate buffer, 1 M EDTA, pH 8. Free sulfhydryl groups of Cys react with Ellman's reagent to yield a mixed disulfide and TNB^{2-} ion, which has a characteristic yellow colour.

The reaction between iodoacetyl-functionalized agarose and free sulfhydryl groups of peptides Cys residues is illustrated in Figure 5.6. The reaction was carried out in 50 mM Tris, 5 mM EDTA-Na, pH 8.5. The pH of the solution should be kept between 7.5 and 9, in order to avoid unwanted side reactions with amine functional groups at higher pH values. EDTA-Na was added to the coupling buffer to prevent disulfide bond formation [37]. As the Ellman's test revealed that 30-40% of the thiol groups were oxidized, a reducing agent (TCEP) was added to the peptide solution prior to resin coupling. TCEP was selected because it does not interfere with the immobilization reaction, since it does not have sulfhydryl groups in its structure [38]. Unreacted sites on iodoacetylated agarose were

blocked using a blocking agent (cysteine, ethanethiol, or β-mercaptoethanol) in order to prevent side reactions, as iodine ion is an excellent leaving group.

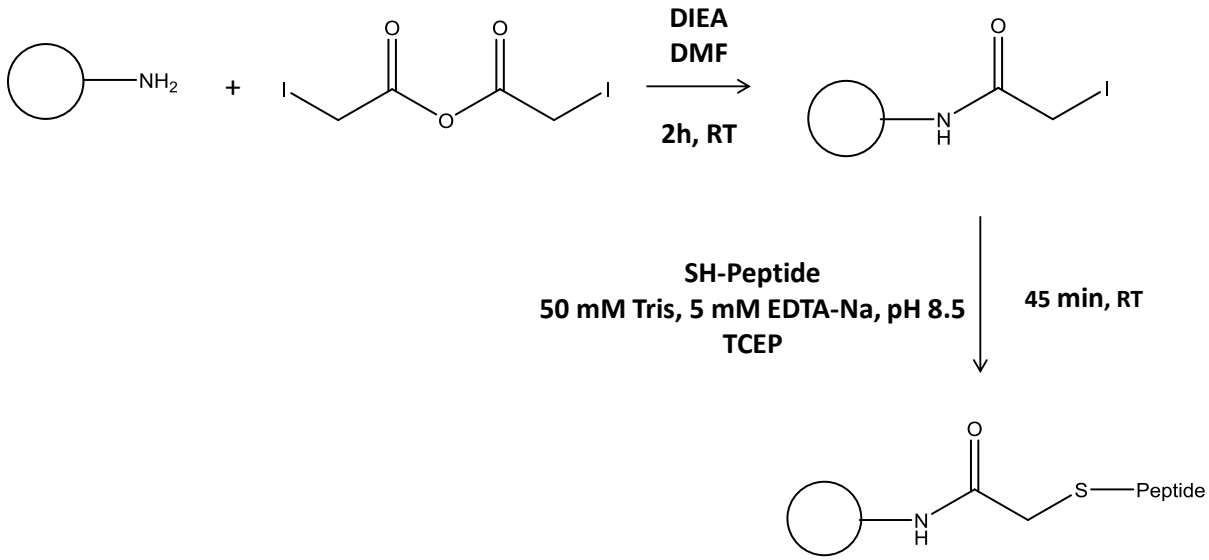


Figure 5.6 – Strategy for the immobilization of peptides through Cys residue. Aminated agarose reacts with iodoacetic anhydride to yield iodoacetylated agarose, which in turn reacts with peptides through their Cys sulfydryl groups.

As TCEP interferes with the BCA quantification method, unbound Cyclic-M0 and Cyclic-M1 were quantified by absorbance at 280 nm in order to determine the yield of immobilization, which was 60% for both peptides.

Quantification of these peptides at 280 nm was possible due to the presence of Pro residues embedded along the peptide. Although Pro residue itself presents a low molar extinction coefficient, it displays increased absorbance when it is part of a peptide chain (except at the N-terminus), probably due to the cyclic nature of the three-carbon side chain to the nitrogen of the peptide backbone. In addition, the presence of two Pro residues together seems to further enhance this effect [39]. Peptides could not be quantified at 205 nm due to the interference of EDTA-Na. At 280 nm neither EDTA-Na (< 30 mM) nor TCEP interfere with peptide quantification [40-42].

A different strategy was followed for the immobilization of Cyclic-M2. The carboxylic group of the unique Glu side chain of Cyclic-M2 was activated by reaction with EDC and NHS, yielding an amine-reactive ester, which will promptly react with aminated agarose (Figure 5.7). This reaction should be carried out in amine- and carboxyl-free buffers. Unbound Cyclic-M2 was quantified using the BCA assay to determine the yield of the immobilization reaction, which was 33%.

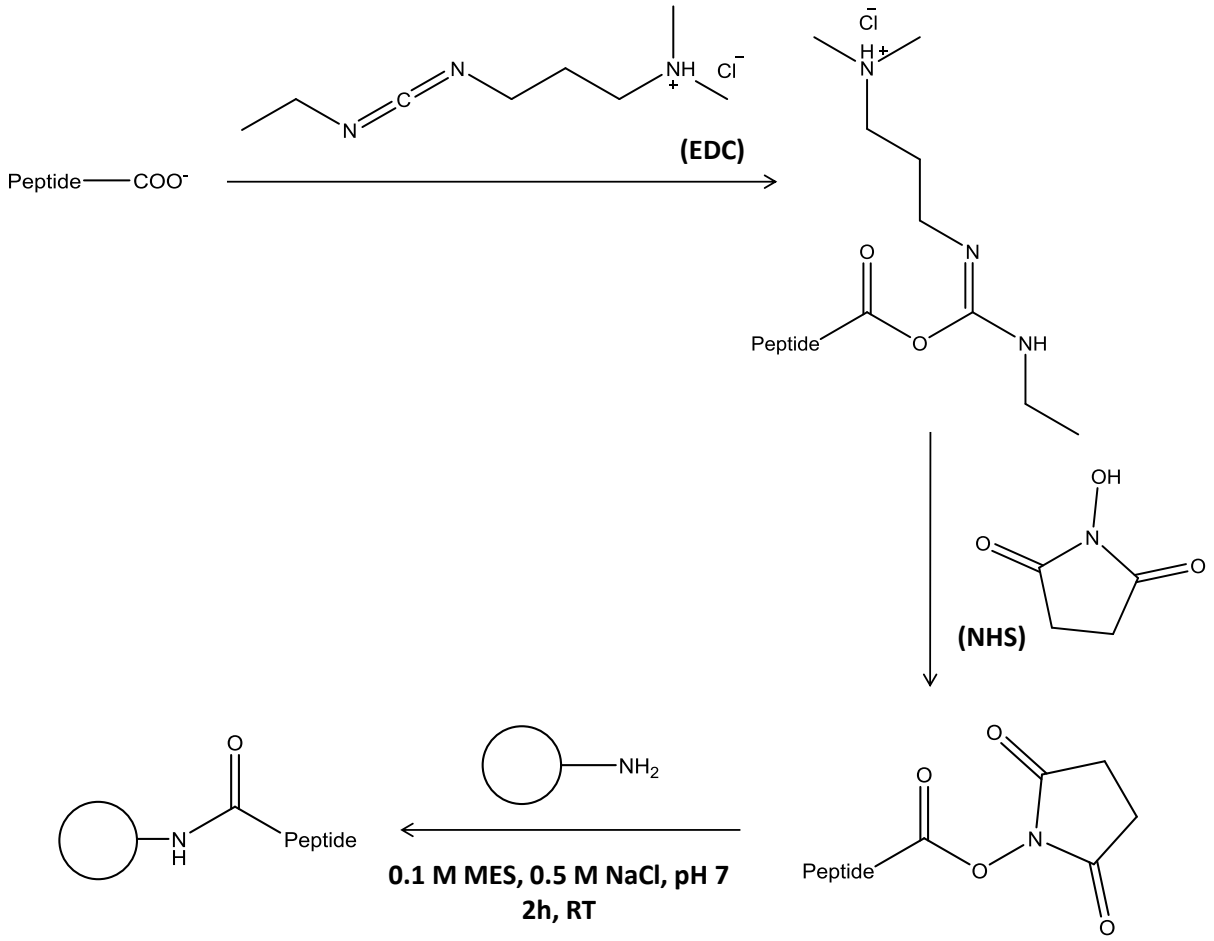


Figure 5.7 – Immobilization of peptides through their carboxylic groups onto aminated agarose. Activation of carboxylic groups of peptides with EDC and NHS yields an amine-reactive ester, which reacts with aminated agarose at pH 7.

5.3.4. Screening β-hairpin cyclic peptides for binding to GK14 and GK14-P peptides

Cyclic-M0 and Cyclic-M1 were screened against their phosphorylated target GK14-P and its non-phosphorylated version at acid and basic pH. Control experiments were also carried out using the blocked agarose solid support (iodoacetylated agarose reacted with cysteine, ethanethiol or β-mercaptoethanol).

Figure 5.8 shows that both Cyclic-M0 and Cyclic-M1 peptides bind the phosphorylated target GK14-P at pH 3. Cyclic-M0 was not expected to bind, as it is a negative control. However, all blocked solid support controls bind exactly the same amount of target peptide than Cyclic-M0 and Cyclic-M1. This means that the problem may not be arising from the lack of specificity of the cyclic β-hairpin towards the target, but from the non-specific binding of the target to the blocked resin.

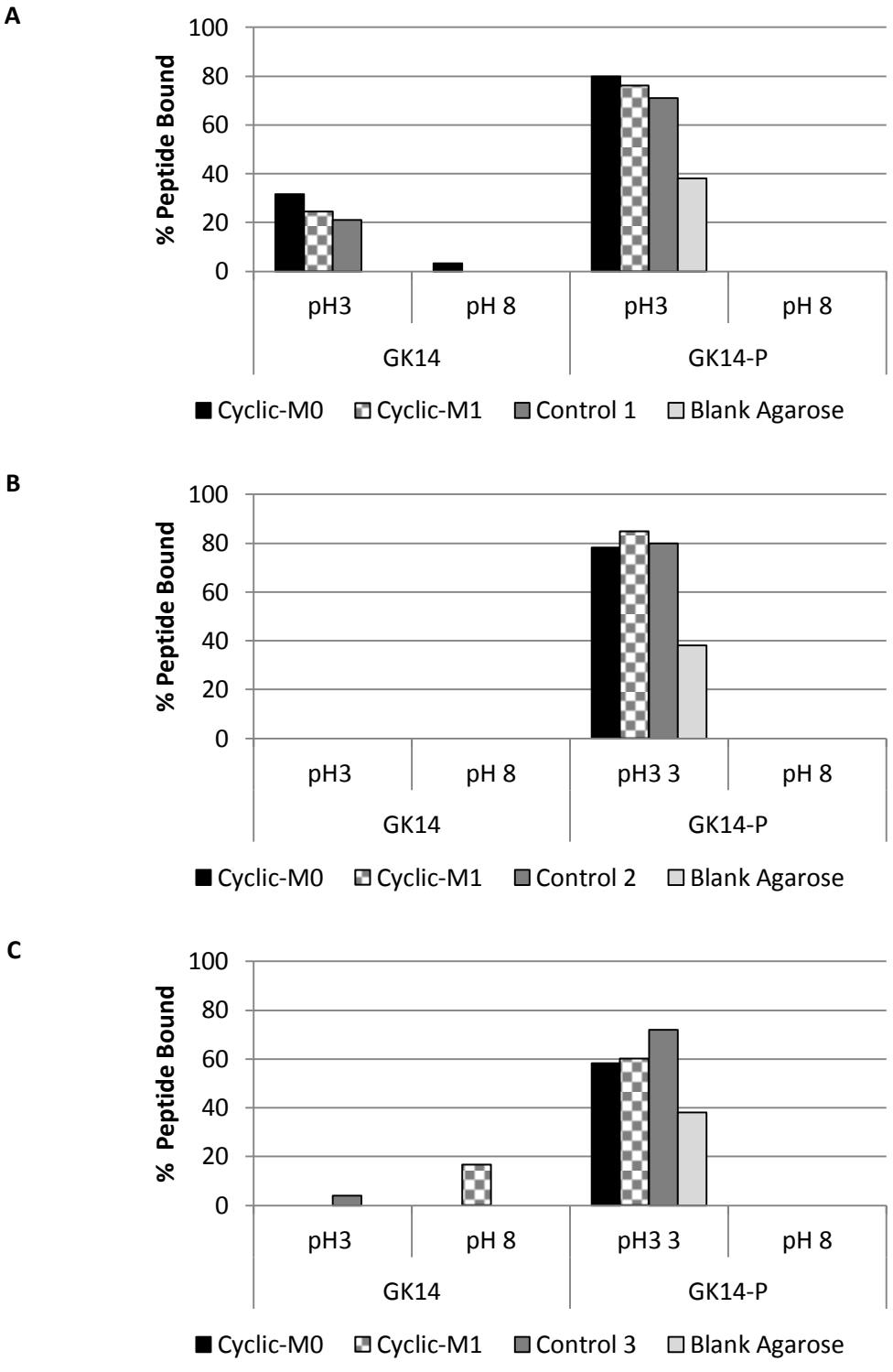


Figure 5.8 – Screening of cyclic peptide-immobilized agarose against GK14 and GK14-P after blocking unreacted sites with: A) ethanethiol; B) β-mercaptoethanol; and C) Cys. Screenings were performed by adding 0.25 mL of peptide in binding buffer (0.5 mg/ mL) to each well of a 96-well filter plate containing 0.25 g of resin. The unbound peptides from flow-through and washes were collected in 96-well microplates and quantified using the BCA assay. Binding buffers were: 50 mM Gly pH 3 and 50 mM HEPES pH 8. Controls correspond to iodoactelyated agarose reacted with A) ethanethiol (Control 1); B) β-mercaptoethanol (Control 2); and C) Cys (Control 3). Blank agarose was also used as control in each case.

Cyclic-M2, which was immobilized through carbodiimide chemistry, was also screened against GK14 and G14-P. The screening was performed at pH 8, because of the non-specific binding observed for agarose (control) at acidic pH. At pH 3 agarose binds approximately 38% to GK14-P, while no binding is observed at pH 8. The non-phosphorylated peptide does not bind to agarose at any of the pH values tested (Figure 5.8).

Cyclic-M2 bound approximately 0.54±0.04 mg GK14-P/g Cyclic-M2-functionalized resin. It was highly selective toward GK14-P, presenting a phosphopeptide mole fraction (defined in §Equation 2.2) of 97%. GK14-P was successfully eluted from the resin using 50 mM HEPES, 250 mM NaCl pH 8, with up to 61±25% recovery.

Immobilization yields and binding capacity for GK14-P of Cyclic-M0, Cyclic-M1, and Cyclic-M2 are presented in Table 5.3.

Table 5.3 – Comparison between Cyclic-M0, Cyclic-M1, and Cyclic-M2 in terms of immobilizations yields and binding capacity for GK14-P (N=2).

Peptide ID	Yield of immobilization (%)[a]	µmol Immobilized cyclic peptide/ g moist resin	Binding at pH 8 (mg GK14-P/g moist resin)	GK14-P mole fraction (%)	Elution (%)
Cyclic-M0	60	48.51±0.48	0	0	-
Cyclic-M1	60	47.26±0.76	0	0	-
Cyclic-M2	33	32.38±0.08	0.54±0.04	97±5	61±25

[a] Yield of immobilization relative to the amount of functional groups in the solid support.

The yield of immobilization is lower for Cyclic-M2, probably due to peptide cross-linking, since it is plausible that some of the activated carboxyl groups of the peptide might react with amine groups of Dap and Arg amino acids, while others will react with aminated agarose. In terms of binding capacity at pH 8, Cyclic-M0 and Cyclic-M1 do not bind to the target, while Cyclic-M2 is highly selective for GK14-P. Cyclic-M2 is illustrated in Figure 5.9.

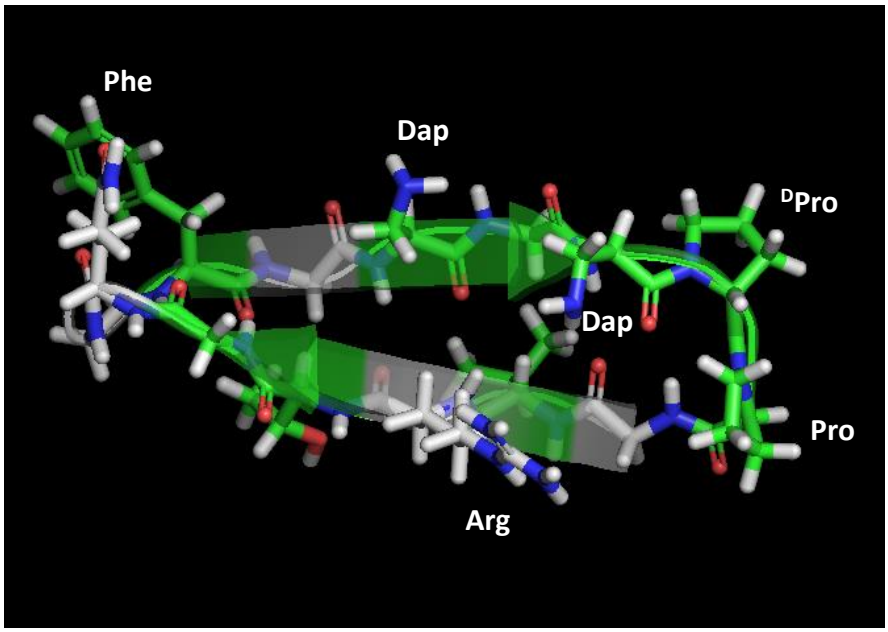


Figure 5.9 – Illustration of Cyclic-M2 sketched in PyMOL Molecular Graphics System v.1.3 (Schrödinger). The diproline unit and the amino acids involved in molecular recognition are highlighted in 3 letter codes.

To better understand the interactions between Cyclic-M2 and GK14-P, both peptides were sketched in PyMOL Molecular Graphics System v.1.3 (Schrödinger) and their properties determined in MarvinSketch 4.1.13 software (ChemAxon) (Table 5.4).

Table 5.4 – Properties of Cyclic-M2 and GK14-P determined using pI and logD plugins of MarvinSketch 4.1.13 software (ChemAxon). pI – Isoelectric point.

	pI	Net Charge at pH 8	logD at pH 8
Cyclic-M2	12.55	2.09	-15.04
GK14-P	3.77	-2.90	-14.30

pI: isoelectric point.

At pH 8 the peptides have opposite net charges, which means they most likely establish electrostatic interactions. Only three residues are ionisable in Cyclic-M2: Dap at position 5 (p*K*a=7.79), Dap at position 7 (p*K*a=8.39), and Arg at position 12 (p*K*a=11.85). GK14-P has five charged functional groups at pH 8: one negative charge on Asp side chain; two negative charges on the phosphate group of pSer; one positive charge on Lys side chain; and one negative charge on the carboxyl terminal group (Figure 5.10). The fact that Cyclic-M2 selectively binds GK14-P but not GK14 indicates that the presence of the phosphate group is vital for the molecular recognition of the peptide.

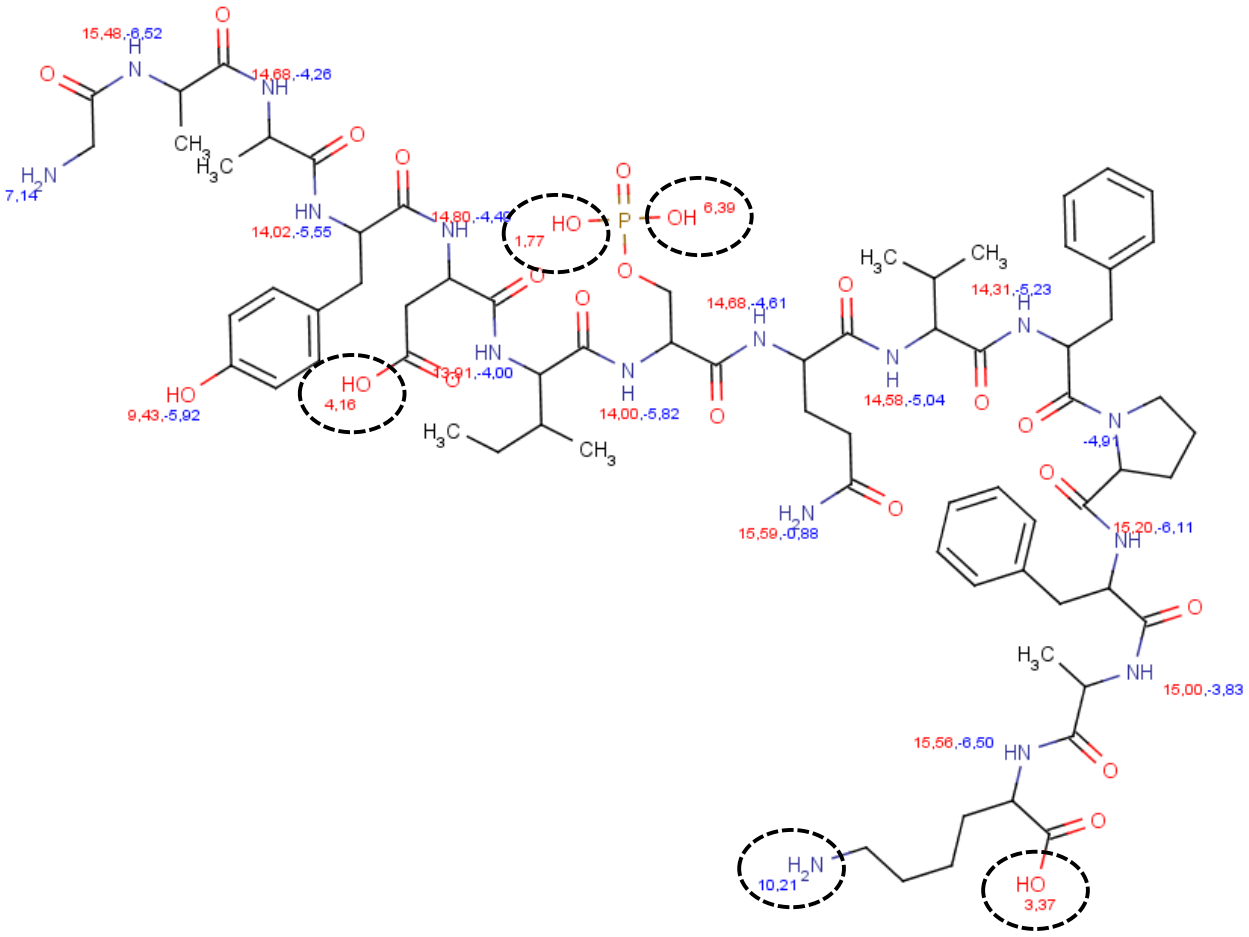


Figure 5.10 – p*Ka* values of the functional groups of GK14-P. The charged amino acids at pH 8 are highlighted with dashed lines. Values were determined using the pKa Plugin from MarvinSketch (ChemAxon).

Negative logD values indicate that both peptides are hydrophilic at pH 8 (Table 5.4). However, this does not exclude the establishment of hydrophobic interactions between the non-polar functional groups of the peptides, as it was observed in the molecular docking studies.

5.4. Conclusions

Three cyclic β-hairpin structures mimicking the BRCA1-BRCT domain were successfully synthesized by chemical methods. Linear precursors were cyclized by two different methodologies, one based on the stepwise addition of the peptide and HATU, and another on the use of PyBOP as coupling reagent and higher amounts of solvents. The latter revealed to be more efficient in this work. Either a Cys or a Glu residue was introduced into the β-hairpins structure in order to facilitate their immobilization

onto an agarose resin after cyclization. Peptides containing Cys (Cyclic-M0 and Cyclic-M1) were attached to iodoacetylated agarose with immobilization yields of 60%, but the need to block the unreacted iodoacetyl groups renders this methodology less effective, because the target peptides (GK14 and GK14-P) bind non-specifically to the solid support. Therefore, the ability of Cyclic-M0 and Cyclic-M1 to specifically recognize GK14-P remains unknown. Cyclic-M2 contained a Glu residue and was therefore immobilized using carbodiimide chemistry, presenting an immobilization yield of 33%. Cyclic-M2 was highly selective at pH 8 towards the phosphorylated peptide GK14-P, which subsequently was successfully eluted from the resin under mild conditions.

5.5. Bibliography

1. Ko, E., J. Liu, and K. Burgess, *Minimalist and universal peptidomimetics.* Chemical Society Reviews, 2011. **40**(8): p. 4411-4421.
2. Wu, Y.-D. and S. Gellman, *Peptidomimetics.* Accounts of chemical research, 2008. **41**(10): p. 1231-1232.
3. Kemp, D.S., *Peptidomimetics and the template approach to nucleation of β-sheets and α-helices in peptides.* Trends in Biotechnology, 1990. **8**: p. 249-255.
4. Grauer, A. and B. Koenig, *Peptidomimetics–A versatile route to biologically active compounds.* European Journal of Organic Chemistry, 2009. **2009**(30): p. 5099-5111.
5. Joo, S.H., *Cyclic Peptides as Therapeutic Agents and Biochemical Tools.* Biomolecules & therapeutics, 2012. **20**(1): p. 19.
6. Joo, S.H., Q. Xiao, Y. Ling, B. Gopishetty, and D. Pei, *High-throughput sequence determination of cyclic peptide library members by partial Edman degradation/mass spectrometry.* Journal of the American Chemical Society, 2006. **128**(39): p. 13000-13009.
7. Millward, S.W., T.T. Takahashi, and R.W. Roberts, *A general route for post-translational cyclization of mRNA display libraries.* Journal of the American Chemical Society, 2005. **127**(41): p. 14142-14143.
8. Hamzeh-Mivehroud, M., A.A. Alizadeh, M.B. Morris, W. Bret Church, and S. Dastmalchi, *Phage display as a technology delivering on the promise of peptide drug discovery.* Drug Discovery Today, 2013. **18**(23): p. 1144-1157.
9. Wang, H. and R. Liu, *Advantages of mRNA display selections over other selection techniques for investigation of protein-protein interactions.* Expert Reviews in Proteomics 2011. **8**(3): p. 335-346.
10. Takahashi, T.T., R.J. Austin, and R.W. Roberts, *mRNA display: ligand discovery, interaction analysis and beyond.* Trends in Biochemical Sciences, 2003. **28**(3): p. 159-165.
11. Evans, T.C., D. Martin, R. Kolly, D. Panne, L. Sun, I. Ghosh, L. Chen, J. Benner, X.-Q. Liu, and M.-Q. Xu, *Protein trans-Splicing and Cyclization by a Naturally Split Intein from the dnaE Gene of Synechocystis Species PCC6803.* Journal of Biological Chemistry, 2000. **275**(13): p. 9091-9094.
12. Perler, F.B., *Protein splicing of inteins and hedgehog autoproteolysis: structure, function, and evolution.* Cell, 1998. **92**(1): p. 1-4.
13. Scott, C.P., E. Abel-Santos, M. Wall, D.C. Wahnon, and S.J. Benkovic, *Production of cyclic peptides and proteins in vivo.* Proceedings of the National Academy of Sciences, 1999. **96**(24): p. 13638-13643.

14. Tavassoli, A. and S.J. Benkovic, *Split-intein mediated circular ligation used in the synthesis of cyclic peptide libraries in E. coli.* Nature Protocols, 2007. **2**(5): p. 1126-1133.
15. Geysen, H.M., R.H. Meloen, and S.J. Barteling, *Use of peptide synthesis to probe viral antigens for epitopes to a resolution of a single amino acid.* Proceedings of the National Academy of Sciences, 1984. **81**(13): p. 3998-4002.
16. Vetter, S. and Z. Zhang, *Combinatorial chemistry and peptide library methods to characterize protein phosphatases.* Methods in Enzymology, 2003. **366**: p. 260-282.
17. Lam, K.S., S.E. Salmon, E.M. Hersh, V.J. Hruby, W.M. Kazmierski, and R.J. Knapp, *A new type of synthetic peptide library for identifying ligand-binding activity.* Nature, 1991. **354**(6348): p. 82-84.
18. Kwon, Y.-U. and T. Kodadek, *Encoded combinatorial libraries for the construction of cyclic peptoid microarrays.* Chem. Commun., 2008(44): p. 5704-5706.
19. Robinson, J.A., *β-Hairpin peptidomimetics: design, structures and biological activities.* Accounts of Chemical Research, 2008. **41**(10): p. 1278-1288.
20. Hughes, R.M. and M.L. Waters, *Model systems for β-hairpins and β-sheets.* Current Opinion in Structural Biology, 2006. **16**: p. 514-524.
21. Pastor, M.T. and E. Pérez-Payá, *Combinatorial chemistry of β-hairpins.* Molecular Diversity, 2003. **6**(2): p. 149-155.
22. Blanco, F., M. Ramirez-Alvarado, and L. Serrano, *Formation and stability of p-hairpin structures in polypeptides.* Current Opinion in Structural Biology, 1998. **8**: p. 107-111.
23. Lewandowska, A., S. Ołdziej, A. Liwo, and H.A. Scheraga, *β-hairpin-forming peptides; models of early stages of protein folding.* Biophysical Chemistry, 2010. **151**(1): p. 1-9.
24. Makwana, K.M., S. Raghothama, and R. Mahalakshmi, *Stabilizing effect of electrostatic vs. aromatic interactions in diproline nucleated peptide β-hairpins.* Physical Chemistry Chemical Physics, 2013. **15**(37): p. 15321-15324.
25. Danelius, E., U. Brath, and M. Erdélyi, *Insight into β-Hairpin Stability: Interstrand Hydrogen Bonding.* SYNLETT, 2013. **24**: p. 2407-2410.
26. Varedian, M., M. Erdélyi, Å. Persson, and A. Gogoll, *Interplaying factors for the formation of photoswitchable β-hairpins: the advantage of a flexible switch.* Journal of Peptide Science, 2009. **15**(2): p. 107-113.
27. White, C.J. and A.K. Yudin, *Contemporary strategies for peptide macrocyclization.* Nature Chemistry, 2011. **3**(7): p. 509-524.
28. Park, J.H. and M.L. Waters, *Positional effects of click cyclization on β-hairpin structure, stability, and function.* Organic & Biomolecular Chemistry, 2013. **11**(1): p. 69-77.
29. Aravinda, S., U.S. Raghavender, R. Rai, V.V. Harini, N. Shamala, and P. Balaram, *Analysis of designed β-hairpin peptides: molecular conformation and packing in crystals.* Organic & Biomolecular Chemistry, 2013. **11**(25): p. 4220-4231.
30. Gibbs, A.C., T.C. Bjorndahl, R.S. Hodges, and D.S. Wishart, *Probing the structural determinants of type II'β-turn formation in peptides and proteins.* Journal of the American Chemical Society, 2002. **124**(7): p. 1203-1213.
31. Chakraborty, K., P. Shivakumar, S. Raghothama, and R. Varadarajan, *NMR structural analysis of a peptide mimic of the bridging sheet of HIV-1 gp120 in methanol and water.* Biochemical Journal, 2005. **390**: p. 573-581.
32. *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*, Chan, W.C., and White, P.D. (Eds.), Oxford University Press, New York, 2000.
33. Carpino, L.A., A. El-Faham, C.A. Minor, and F. Albericio, *Advantageous applications of azabenzotriazole (triazolopyridine)-based coupling reagents to solid-phase peptide synthesis.* Journal of the Chemical Society, Chemical Communications, 1994(2): p. 201-203.

34. Malesevic, M., U. Strijowski, D. Bächle, and N. Sewald, *An improved method for the solution cyclization of peptides under pseudo-high dilution conditions.* Journal of Biotechnology, 2004. **112**(1): p. 73-77.
35. Zhou, H., J.D. Watts, and R. Aebersold, *A systematic approach to the analysis of protein phosphorylation.* Nature Biotechnology, 2001. **19**(4): p. 375-378.
36. Myc, A., A.K. Patri, and J.R. Baker, *Dendrimer-based BH3 conjugate that targets human carcinoma cells.* Biomacromolecules, 2007. **8**(10): p. 2986-2989.
37. Trivedi, M.V., J.S. Laurence, and T.J. Siahaan, *The role of thiols and disulfides in protein chemical and physical stability.* Current Protein & Peptide Science, 2009. **10**(6): p. 614.
38. Hansen, R.E. and J.R. Winther, *An introduction to methods for analyzing thiols and disulfides: reactions, reagents, and practical considerations.* Analytical Biochemistry, 2009. **394**(2): p. 147-158.
39. Kuipers, B.J. and H. Gruppen, *Prediction of molar extinction coefficients of proteins and peptides using UV absorption of the constituent amino acids at 214 nm to enable quantitative reverse phase high-performance liquid chromatography-mass spectrometry analysis.* Journal of Agricultural and Food Chemistry, 2007. **55**(14): p. 5445-5451.
40. Demeler, B., *Methods for the design and analysis of sedimentation velocity and sedimentation equilibrium experiments with proteins.* Current Protocols in Protein Science, 2010: p. 7.13.1-7.13.24.
41. Stoscheck, C.M. (1990) *[6] Quantitation of protein*, In P.D. Murray (Ed.), Methods in Enzymology, Academic Press, Volume 182, p. 50-68.
42. Noble, J.E. and M.J. Bailey, *Quantitation of protein.* Methods in Enzymology, 2009. **463**: p. 73-95.

CHAPTER 6
Concluding remarks

Two types of synthetic affinity reagents were developed for the specific binding of phosphorylated peptides: small organic ligands and cyclic β-hairpin peptides.

The small synthetic ligands were designed to mimic the interactions occurring in the phosphate-binding pockets of the human phosphoprotein-binding domains. Very recently, Marx and co-workers built 96 combinatorial libraries comprising more than 100,000 phosphorylated and non-phosphorylated synthetic peptides. The idea was to acquire a large number of mass spectrometric data to evaluate the performance of several computational tools (e.g. Mascot) currently employed for peptide identification and phosphorylation-site localization. They have also compared different fragmentation methods (beam-type collision-induced dissociation (HCD) and electron transfer dissociation (ETD)) and characterized the chromatographic profile of the synthesized peptides [1]. While they have focused on the combinatorial synthesis of phosphorylated peptides, we developed combinatorial libraries of phosphopeptide-binders. Both strategies represent complementary tools to facilitate the identification and profiling of clinically relevant phosphorylated proteins and peptides.

The combinatorial libraries were based on three chemical reactions: (i) nucleophilic substitutions on a triazine ring; (ii) Ugi multicomponent reaction; and (iii) tandem Petasis-Ugi multicomponent reaction. The feasibility of these reactions on agarose was confirmed by Inductively Coupled Plasma - Atomic Emission Spectroscopy (ICP-AES) and fluorescence microscopy (Table 6.1). Petasis-Ugi scaffold presented the lowest reaction yield; therefore, synthesis conditions should be further optimized, by testing different solvents, temperatures, and reaction times. To our knowledge, this is the first time that a Petasis-Ugi scaffold has been reported for affinity chromatography.

Table 6.1 – Reaction yields of triazine, Ugi, and Petasis-Ugi reactions on agarose determined from ICP-AES analysis.

	Scaffolds		
	Triazine (Ligand A10A10)	Ugi (Ligand A10C2)	Petasis-Ugi (D3bA10B2)
Reaction Yields (%)	100	94	44

Chapter 6 Concluding remarks

A variety of amines, aldehydes, and carboxylic acids were used as building blocks for the introduction of chemical diversity in the scaffold molecules, mimicking key amino acid residues involved in phosphate recognition. After several steps of selection, two lead ligands were identified – A8A3 and A8C2, based on triazine and Ugi reactions, respectively. Interestingly, both of these ligands present a chemical group mimicking a Histidine amino acid residue, confirming a role for the imidazole group on phosphate coordination, most likely trough the establishment of both hydrophobic and electrostatic interactions.

A8A3 and A8C2 were screened against both mono- and multi-phosphorylated peptides in a wide pH range (3-11) and NaCl concentrations (0-1M), and both ligands presented higher binding capacities and higher selectivity at pH 3. The presence of low concentrations of NaCl enhanced the selectivity for multi-phosphorylated peptides, but not for mono-phosphorylated species. Both acidic and basic conditions can be used to elute phosphopeptides from these ligands. In particular, the addition of H_3PO_4 to the elution solution allows for the recovery of both mono- and multi-phosphorylated peptides from both ligands.

Partition equilibrium experiments indicated that weak affinity interactions govern ligand-peptide binding, providing improved kinetics, high-resolution and rapid throughput.

Ligands A8A3 and A8C2 were further immobilized onto iron oxide magnetic nanoparticles coated with silica and dextran (MNP-Si-Si-Dex). Ligands immobilized on MNP-Si-Si-Dex presented enhanced binding capacities (10-fold higher) and lower selectivity (especially for the mono-phosphorylated species) towards the phosphopeptide targets than the same ligands immobilized on agarose.

MNP-Si-Si-Dex-A8A3 and MNP-Si-Si-Dex-A8C2 were further screened against a tryptic semi-complex mixture of two phosphorylated proteins - α-casein and β-casein - and one non-phosphorylated protein - bovine serum albumin. The mass spectrum of MNP-Si-Si-Dex-A8C2 presented a low signal-to-noise ratio and no peptides could be identified (neither phosphorylated nor non-phosphorylated). The mass spectrum of MNP-Si-Si-Dex-A8A3 was dominated by the presence of peaks correspondent to phosphorylated peptides. Nine phosphorylated peptides were identified from a 6 pmol semi-complex sample, comprising both mono- and multi-phosphorylated peptides. A Ti^{4+}-IMAC approach was tested in parallel and yielded equivalent results. Future work includes determining the sensitivity of our method, by testing samples with lower concentrations and decreasing the ratio between phosphopeptides and non-phosphopeptides in the original tryptic sample.

A different line of study was followed for the synthesis of cyclic β-hairpin peptidomimetics. Instead of performing broad structural studies of the ten human PDBs and synthesizing ligands in a combinatorial manner, we have focused on a specific domain known to be involved in breast and ovarian cancers: the Brca1 C-terminal (BRCT) domain. Based on the crystallographic structure of the

BRCA1-BRCT domain in complex with its phosphorylated target Gly-Ala-Ala-Tyr-Asp-Ile-pSer-Gln-Val-Phe-Pro-Phe-Ala-Lys (GK14-P), a structure-based design approach has been employed for the development of cyclic β-hairpin peptides. These peptides act as miniaturized versions of the BRCT domain, mimicking specific regions of phosphate recognition. A lead candidate with the sequence cyclo(Glu-Gly-Phe-Gly-Dap-Gly-Dap-^{D}Pro-Pro-Gly-Val-Arg-Thr-Gly) – Cyclic-M2 - has been identified, presenting high binding capacity and selectivity towards GK14-P.

This concept can be applied to other PDBs and drive the development of highly specific diagnostic tools. For example, SH2 domains have been used in large-scale far-western analysis and reverse-phase protein arrays to profile phosphoprotein and phosphopeptide binding [2]. The employment of miniaturized versions of PBDs would avoid the laborious synthesis of whole domains, while helping to elucidate the molecular mechanisms behind a variety of cellular processes. Conversely, small non-peptide SH2 ligands have been developed as tools for the discovery of novel drugs by blocking SH2-phosphoprotein interactions [3].

It should be noted that both A8A3 and A8C2 present their highest selectivity towards a phosphopeptide with the sequence pSer-Gln-Val-Phe-Pro-Trp (SW6-P), which comprises a consensus sequence known to specifically bind to BRCT domain. The three peptidomimetic scaffolds immobilized on agarose were compared in terms of binding capacity and selectivity towards their target peptides (Table 6.2). Cyclic-M2 presents a lower binding capacity, but it employs mild binding and elution conditions, which might be a competitive advantage, especially for the enrichment of phosphoproteins in non-denaturating conditions. The commercially available Phos-tag^{TM} Agarose reagent for phosphopeptide separation is reported to have 3-5 µmol Phos-tag/ mL-gel [4]. However, to our knowledge, there is no data available referring to the amount of phosphopeptides that actually bind to the support. Results are usually evaluated by SDS-PAGE and Western blotting [5, 6].

Table 6.2 – Binding capacity (µmol target peptide/ g moist agarose) and phosphopeptide mole fraction (%) of ligands A8A3, A8C2, and Cyclic-M2 immobilized on agarose towards their respective target peptides. Screenings were performed at optimal binding conditions in each case. A8A3 and A8C2 were screened against SW6-P at pH 4, and Cyclic-M2 was screened against GK14-P at pH 8. Peptides were quantified by fluorescence microscopy in case of A8A3 and A8C2, and by BCA assay in case of Cyclic-M2.

	Peptidomimetics		
	A8A3	A8C2	Cyclic-M2
Phosporylated peptide bound (µmol target peptide/ g moist agarose)	0.9	1	0.3
Phosphopeptide mole fraction (%)	100	97	97
Target peptide	*Peptide ID: SW6-P* **pS-Q-V-F**-P-W		*Peptide ID: GK14-P* G-A-A-Y-D-I-**pS-Q-V-F**-P-F-A-K

Both strategies described in this work, either based on the combinatorial synthesis of small peptidomimetics or on the rational design of miniaturized versions of PBDs, present novel alternatives to conventional antibody-based and metal-chelating methodologies. All these methodologies do not necessary mutually exclude each other. On the contrary, in our view, the high complexity of phosphorylation events and the difficulties to isolate clinically-relevant phosphorylated species creates a need to build more comprehensive, integrative, and complementary approaches for phosphopeptide enrichment.

Bibliography

1. Marx, H., S. Lemeer, J.E. Schliep, L. Matheron, S. Mohammed, J. Cox, M. Mann, A.J. Heck, and B. Kuster, *A large synthetic peptide and phosphopeptide reference library for mass spectrometry-based proteomics.* Nature Biotechnology, 2013. **31**(6): p. 557-564.
2. Machida, K., C.M. Thompson, K. Dierck, K. Jablonowski, S. Kärkkäinen, B. Liu, H. Zhang, P.D. Nash, D.K. Newman, and P. Nollau, *High-throughput phosphotyrosine profiling using SH2 domains.* Molecular Cell, 2007. **26**(6): p. 899-915.
3. Plummer, M.S., D.R. Holland, A. Shahripour, E.A. Lunney, J.H. Fergus, J.S. Marks, P. McConnell, W.T. Mueller, and T.K. Sawyer, *Design, synthesis, and cocrystal structure of a nonpeptide Src SH2 domain ligand.* Journal of Medicinal Chemistry, 1997. **40**(23): p. 3719-3725.
4. *Phos-tag.com*. Available from: http://www.phos-tag.com/, Access Date: 17 May 2014.
5. Kinoshita, E., A. Yamada, H. Takeda, E. Kinoshita-Kikuta, and T. Koike, *Novel immobilized zinc (II) affinity chromatography for phosphopeptides and phosphorylated proteins.* Journal of Separation Science, 2005. **28**(2): p. 155-162.
6. Kinoshita-Kikuta, E., E. Kinoshita, A. Yamada, M. Endo, and T. Koike, *Enrichment of phosphorylated proteins from cell lysate using a novel phosphate-affinity chromatography at physiological pH.* Proteomics, 2006. **6**(19): p. 5088-5095.